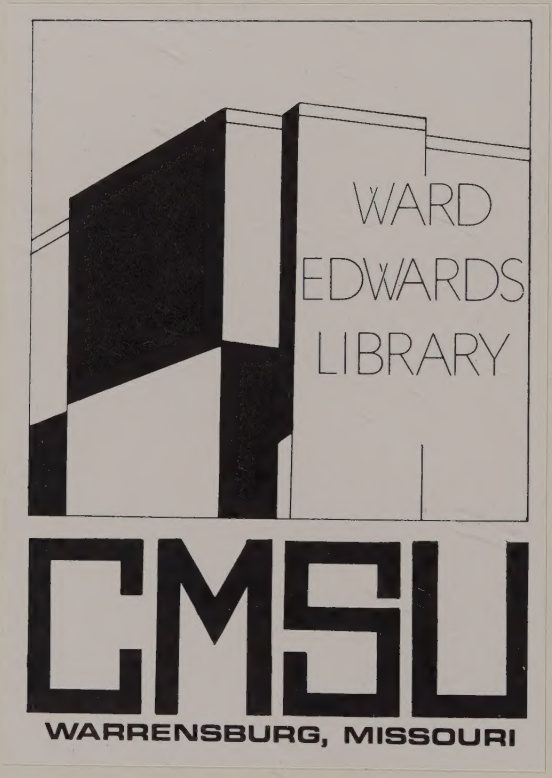
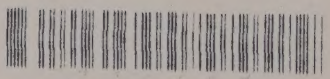
THERMODYNAMICS AND STATISTICAL PHYSICS

Don C. Kelly





WARD
EDWARDS
LIBRARY

CMSU

WARRENSBURG, MISSOURI

0334844

QC 311.15K44
THERMODYNAMICS AND ST
ATISTICAL PHY

KELLY DON C.

WITHDRAWN BY
JAMES C. KIRKPATRICK LIBRARY
TECHNICAL SERVICES

Thermodynamics and Statistical Physics

Thermodynamics and Statistical Physics

AN ELEMENTARY TREATMENT
WITH CONTEMPORARY APPLICATIONS

DON C. KELLY

Miami University
Oxford, Ohio

CENTRAL MISSOURI
STATE UNIVERSITY
Warrensburg,
Missouri

ACADEMIC PRESS New York and London

COPYRIGHT © 1973, BY ACADEMIC PRESS, INC.
ALL RIGHTS RESERVED.
NO PART OF THIS PUBLICATION MAY BE REPRODUCED OR TRANSMITTED IN ANY FORM OR BY ANY MEANS, ELECTRONIC OR MECHANICAL, INCLUDING PHOTOCOPY, RECORDING, OR ANY INFORMATION STORAGE AND RETRIEVAL SYSTEM, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHER.

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 72-88354

PRINTED IN THE UNITED STATES OF AMERICA

QC
311.15
K44

To Jane

334844

Contents

Preface xiii

Overview 1

Chapter 1 **A Framework for Thermodynamics**

1.1 Preliminary Notions 3
1.2 States—Mechanical and Thermodynamic 6
1.3 Thermodynamic Equilibrium and Temperature 8
1.4 The Zeroth Law of Thermodynamics 9
1.5 Pressure—An Important Thermodynamic Variable 10
1.6 Temperature and Thermometry 15
References 21

Chapter 2 **Mathematical Preliminaries**

2.1 Differentials 23
2.2 Partial Derivatives 26
2.3 Two Important Partial Derivative Relations 29
2.4 Using Partial Derivatives in Thermodynamics 31
References 36

Chapter 3 **The First Law of Thermodynamics**

3.1 The Many Faces of Energy 37
3.2 Calorimetric Definition of Heat 43
3.3 Work in Thermodynamics 46
3.4 Internal Energy 48
3.5 The First Law of Thermodynamics 49
References 53

Chapter 4 **Thermodynamic Processes and Thermodynamic Work**

4.1	Thermodynamic Processes	54
4.2	Quasi-Static Processes	55
4.3	Reversible Processes and Thermodynamic Work	58
4.4	Magnetic Work	65
4.5	Heat Capacities and Internal Energy	69
	References	76

Chapter 5 **The Second Law of Thermodynamics**

5.1	Introduction	77
5.2	The Kelvin–Clausius Formulations of the Second Law of Thermodynamics	78
5.3	The Carnot Cycle	81
5.4	Thermodynamic Efficiency	83
5.5	Carnot's Theorem and the Thermodynamic Temperature Scale	87
5.6	The Clausius–Clapeyron Equation	91
5.7	Maxwell's Demon	97
	References	100

Chapter 6 **Entropy**

6.1	Entropy and the Carnot Cycle	102
6.2	Entropy—Examples and Insights	107
6.3	The Ideal Gas, Free Expansion, and Gibbs' Paradox	111
6.4	*T–S* Diagrams	116
6.5	Entropy Formulation of the Second Law of Thermodynamics	119
6.6	Entropy and Universal Accidents	120
	References	122

Chapter 7 **The Third Law of Thermodynamics**

7.1	Entropy, Absolute Zero, and the Third Law	123
7.2	The Third Law in Action	125
7.3	Exceptions and Restatement	129
7.4	The Unattainability of Absolute Zero	133
	References	135

Chapter 8 **Thermodynamic Potentials**

8.1	Introduction	137
8.2	The Maxwell Relations	141
8.3	Helmholtz Free Energy	145
8.4	Enthalpy and Thermochemistry	147
8.5	Gibbs Free Energy	151

8.6 Chemical Potential 154
References 159

Chapter 9 **Phase Changes and Low-Temperature Physics**

9.1 Phase Changes and Phase Diagrams 161
9.2 Liquefaction and Cryogenics 170
9.3 Magnetic Cooling 175
References 182

Chapter 10 **Elementary Statistical Concepts**

10.1 Mean Values and Probabilities 184
10.2 Conditional Probabilities 192
10.3 Continuous Variables and Distribution Functions 194
10.4 Free Path Lengths and Their Distribution 198
References 201

Chapter 11 **Gases: Models and Equipartition**

11.1 The Ideal Gas 202
11.2 Ideal Gas Model and Critique 205
11.3 Equation of State 207
11.4 The Equipartition of Energy 211
11.5 Internal Energy and Heat Capacities 216
11.6 Pressure, Energy Density, and Departures from Ideal Gas Behavior 222
References 229

Chapter 12 **Velocity Space Distribution Functions**

12.1 Introductory Ideas 230
12.2 A Distribution Function for Velocity Space 233
12.3 Mean Values 235
12.4 Spherical Coordinate Space and Spherical Velocity Space 236
12.5 The Distribution of Molecular Speeds 240
12.6 Experimental Verification of the Maxwellian Distribution 243
References 245

Chapter 13 **Transport Processes**

13.1 Introduction 247
13.2 Fluxes 248
13.3 Molecular Effusion 251
13.4 Thermal Conduction 253
References 257

Chapter 14 **Mathematics for Classical Statistics**

14.1	Probability Versus Number of Ways	258
14.2	Factorials	260
14.3	Distinguishable Objects and Distributions	260
14.4	Saddle-Point Integration	265
	References	272

Chapter 15 **Classical Statistical Mechanics**

15.1	The Rationale of Statistical Mechanics	273
15.2	Phase Spaces	278
15.3	Phase Spaces and Hotels: An Analogy	281
15.4	Microstates and Macrostates	283
15.5	Thermodynamic Probability	284
15.6	The Most Probable Macrostate and Thermodynamic Equilibrium	286
15.7	The Boltzmann Distribution	287
15.8	The Partition Function for an Ideal Gas	292
15.9	The Statistical Interpretation of Entropy	297
15.10	Thermodynamic Properties: Relationship to the Partition Function	301
15.11	Summary	303
	References	304

Chapter 16 **Applications of Statistical Mechanics**

16.1	Microscopic Basis of the Clausius–Clapeyron Equation	305
16.2	Chemical Reactions	309
16.3	The Langevin Theory of Paramagnetism	316
	References	321

Chapter 17 **Quantum Physics and Quantum Statistics**

17.1	Introduction	322
17.2	The Origins of Quantum Theory	323
17.3	Angular Momentum and Statistics	327
17.4	Fermi–Dirac Statistics	329
17.5	Bose–Einstein Statistics	340
17.6	QSM or CSM?	342
	References	345

Chapter 18 **Applications of Quantum Statistical Mechanics**

18.1	Blackbody Radiation, Big Bang, and the Primeval Fireball	347
18.2	Stimulated Emission: The Laser	357
18.3	White Dwarfs, Neutron Stars, and Pulsars	365
18.4	^{3}He–^{4}He Dilution Refrigerators	376

18.5 Superconductivity 384
18.6 The Ultimate Temperature 399
References 404

Appendix A **Gaussian Integrals** 409

Appendix B **The Universal Character of β** 414

Subject Index 417

Preface

This book evolved from the author's experience in teaching a course in thermodynamics and statistical physics at Miami University to undergraduate students majoring in physics or engineering. Their previous exposure to the subject was a very brief survey in the freshman course. All portions of the text have been class tested in mimeographed note form—many of them several times.

The primary object of the book is a modest one. It is to help the student arrive at the point where his conceptions of the fundamental ideas are not only correct and precise, but where he feels comfortable with them as well. A lasting concept must seem "handmade."

The scope of the material is limited. The kinetic theory treatment stops short of the Boltzmann equation and its ramifications. The chapters on statistical mechanics make no reference to ensembles. The Boltzmann equation and ensemble theory are not elementary topics to the typical student.

The mathematical level grows in sophistication as the book progresses. Chapters 2, 10, and 14 serve to introduce and/or review various mathematical techniques.

Problems are interspersed within the text so that the student can immediately see how strong a grasp of the material he has. Numerous examples, serving various ends, are included.

References at the end of each chapter range from popular expositions suited to student use to journal articles which the instructor can digest for use as supplementary material.

My thanks are extended to the many students who made helpful suggestions. The comments of Mr. Kent Eschenberg were particularly helpful. The author is especially grateful to Professor George B. Arfken for reading the manuscript and for offering important suggestions. Above all I wish to thank my wife Jane—who was an inspiration in the writing of the book.

Overview

Thermodynamics and statistical physics deal with the properties of bulk matter under circumstances where the notions of temperature and heat cannot be ignored. In their present form the laws of thermodynamics and related principles comprise the essence of 150 years of experimentation and theoretical interpretation. Largely because of its empirical nature thermodynamics remains a formidable subject, especially at an elementary level. In part, the difficulty stems from the almost universal applicability of these laws. The great generality of the laws of thermodynamics requires that they be independent of the detailed workings of any particular physical system. Instead, they are based on universal principles such as the conservation of energy (first law of thermodynamics). The definitions of thermodynamic variables, though operational[†] and straightforward, may not convey much of a feel for the quantity so defined. This stems from the fact that such variables are nonmechanical. Historically, thermodynamics did not evolve on the basis of any atomic or mechanical model of matter which lets you see how it works.

There are four laws of thermodynamics, labeled zeroth, first, second, and third. Chronologically, only the third law is correctly numbered. The second law was formulated in 1824. The first law was stated about 20 years later. The third and zeroth laws were put forth in the 20th century.

[†] An operational definition defines a quantity by describing how to measure it.

The zeroth law relates the ideas of temperature equality and thermodynamic equilibrium, without suggesting what temperature measures. The zeroth law formalizes an important experimental fact, namely, that the mutual thermodynamic equilibrium of two systems demands equality of but one property—the property we call "temperature."

The first law recognizes heat as a form of energy, and is interpreted as a statement of energy conservation.

The second law places limits on the extent to which heat may be converted into mechanical energy (work) and other useful forms. The second law is linked with the irreversibility of processes which occur spontaneously in nature, such as the flow of heat.

The third law concerns the absolute zero of temperature. One consequence of the third law is that all schemes by which absolute zero is approached suffer a severe case of diminishing returns, rendering this temperature unattainable. The third law has not discouraged low-temperature research, as temperatures within a few millionths of a degree of absolute zero have been achieved.

By contrast, statistical physics is more transparent. One starts with a microscopic model of the physical system. For example, a rarified gas is envisioned as a collection of microscopic particles (atoms or molecules) subject to the laws of mechanics. Physically observable quantities are identified as average values of certain microscopic quantities. For example, we will see that the temperature of a gas is a measure of the average kinetic energy of its molecules. One can follow the transition from the microscopic realm of atoms and molecules to the macroscopic level of bulk matter and thus develop an intuitive feeling for the quantities introduced in the process.

In general, all but the simplest models present enormously complicated mathematical problems. These complexities arise from the many-body nature of the problems. In solids, liquids, and dense gases, any given atom or molecule is most likely to be found being pushed and pulled by many neighboring atoms. Future successes in many areas of statistical physics appear to be tied to the development of improved methods for handling the many-body problem.

If the successes of statistical physics were restricted to a few idealized models, one might seriously question its methods or even its foundations. However, one overpowering result of great generality ensures our confidence in the formalism: It is possible to deduce the laws of thermodynamics by applying the principles of statistical physics. Thus statistical physics can guarantee that a Dewar of liquid helium will obey the laws of thermodynamics. In addition it provides a formalism which, in principle (but not yet fully in practice), allows us to relate the many measurable properties of bulk helium to the microscopic structure of a single helium atom.

Chapter 1

A Framework for Thermodynamics

1.1 PRELIMINARY NOTIONS

"Pick a system" is advice offered freely to physics students. The *system* is the focus of attention—that chunk of the universe whose behavior we strive to understand. The *surroundings* of a system are the remaining portions of the universe which directly affect the behavior of the system.

Favorite mechanical systems include apples which fall vertically under the influence of gravity and blocks which slide down frictionless planes. Mechanics emphasizes a system's response to external forces. A favorite system in thermodynamics and statistical physics is a rarified gas housed in a cylinder fitted with a piston. In thermodynamics we are generally concerned with the effects of energy exchanges between the system and its surroundings.

The primary virtue of systems like the falling apple and sliding block is their remarkable simplicity. By applying the principles of mechanics—Newton's laws of motion—one can predict the behavior of a falling apple. The predicted behavior is in accord with both experiment and common sense—our everyday experience. Success in dealing with such transparent problems strengthens our confidence in the principles of Newtonian mechanics.

The gas-piston system of Figure 1.1c serves the same purpose in thermodynamics; it illustrates the principles at an intuitive level. Be forewarned, however, that your deductions as beginners in thermodynamics often will not seem to be a matter of common sense. In large part the difficulty in acquiring

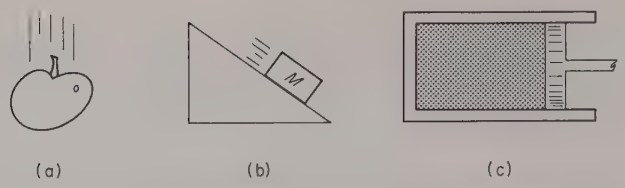


Figure 1.1. Favorite systems in mechanics: (a) apple falling vertically due to gravity, (b) sliding block on a frictionless plane, and in statistical physics, (c) gas–piston system.

a feel for thermodynamics stems from a lack of firsthand experience with the thermal aspects of even the simplest systems, such as gas-filled cylinders.

As a first step toward understanding the behavior of a physical system we must choose quantities which describe the system and the effects of the surroundings. If these physical quantities can be *measured*, either directly or indirectly, they are termed *macroscopic*. For example, if the system is a column of mercury in a thermometer, the density and column length, both measurable, qualify as macroscopic properties. For a confined gas, like that of Figure 1.1c, temperature, density, pressure, and volume are macroscopic variables. The macroscopic quantities used to describe the condition of a thermodynamic system are generally referred to as *thermodynamic variables*.

One may also describe a physical system at the microscopic level. This requires that we first introduce a conceptual model of the system. The model is constructed at the atomic level and employs *microscopic* variables—quantities which are not directly measurable. For example, a gas is envisioned as a collection of an enormous number of atoms which careen about wildly. The mass and kinetic energy of an atom are microscopic variables. The total mass of the gas is measurable, but not the mass of a particular atom.

Thermodynamics is formulated in terms of macroscopic variables. Kinetic theory and statistical mechanics employ microscopic variables in schemes which correlate various features of models with the macroscopic properties of physical systems.

....

1.1.1 Which of the quantities listed below are macroscopic and which are microscopic?

(a) the pressure which drives a piston in an automobile engine,
(b) the force exerted by an atom in a gas when it collides with the container wall,
(c) the wavelength of a water wave,
(d) the speed of a supersonic aircraft,
(e) the speed of a neutron in an atomic nucleus.

....

Thermodynamic variables are further classified according to their dependence on the size of the system (as measured by its mass or volume). A variable is termed *extensive* if its value depends on the size of the system. A variable is referred to as *intensive* if its value is independent of the size of the system. For example, the mass of a system is an extensive property. Its mass density (mass per unit volume) is an intensive quantity.

....

1.1.2 Indicate which of the following are extensive quantities:

(a) the electrical resistance of a length of copper wire,
(b) the electrical resistivity of a length of copper wire,
(c) the kinetic energy of an automobile,
(d) the index of refraction of glass,
(e) the moment of inertia of a flywheel.

....

Thermodynamic systems are classified further with respect to the exchange of matter and energy with their surroundings. In this connection it is necessary to regard heat as a form of energy which is transferred from one region to another by virtue of a temperature difference between them. A more precise and operational definition of heat is presented in Chapter 3.

An *isolated* system can exchange neither matter nor energy with its surroundings. The only candidate for a truly isolated system is the universe. However, it is not certain that all of the laws of thermodynamics apply on a cosmic scale. Despite the fact that an isolated system is an abstraction which cannot be realized in practice, a system may be regarded as isolated whenever matter and energy exchanges do not significantly disturb the processes occurring within the system. For example, a fluid in a container with rigid impermeable walls which are covered with insulation approximates an isolated system. The rigidity of the walls prevents any compressive work being done on the fluid and the insulation minimizes heat transfer. Thus no matter and a minimal amount of energy can be exchanged between the fluid and its surroundings.

A *closed* system is capable of exchanging energy, but not matter, with its surroundings. The walls of a building closely approximate a closed system. They engage in the transfer of thermal energy with the environment without an appreciable exchange of matter.

An *open* system is capable of exchanging both matter and energy with its surroundings. A living animal is a good example of an open system. Matter is exchanged with the surroundings by the acts of eating, breathing, and excreting. Energy exchanges include the absorption and rejection of heat and the performance of mechanical work by body muscles.

In connection with heat flow, we note that extrapolation from experience

suggests the idea of a perfect insulator. The technical name for this wondrous device is *adiabatic wall*. A system enclosed by adiabatic walls is unaffected by thermal changes (as opposed to mechancial changes) in its surroundings, and is said to be "thermally isolated." A system is said to be "in thermal contact with its surroundings" whenever its boundaries do not function as adiabatic walls. Two systems in thermal contact are capable of exchanging heat and are often said to be "separated by diathermic or conducting walls."

Students are often disturbed when the concept of an adiabatic wall is thrust upon them, hailed as fundamental, and then used frequently. They are made uneasy by a feeling that thermodynamics is based on a fantasy. How can thermodynamics deal meaningfully with the real world by introducing an abstraction such as the adiabatic wall; that is, something which cannot be "built"? The notion of an adiabatic wall is valid and useful in thermodynamics in precisely the same way that the idea of a frictionless surface is helpful in mechanics: Both can be approached, that is, almost achieved. Mechanical friction cannot be eliminated completely. Nevertheless it is often possible to reduce frictional forces to the point where they are ignorable by comparison with other forces. The lack of complete thermal isolation is likewise often ignorable.

....

1.1.3 Construct a brief list of physical systems, classifying them as open, closed, or isolated. Be prepared to explain your choices of classification, including the extent to which your entries under closed and isolated are approximations.

....

1.2 STATES—MECHANICAL AND THERMODYNAMIC

Consider a system composed of N particles. The particles might be the atoms comprising your brain or the molecules of the air in some classroom. The mechanical state of the system is fixed by a specification of the position and velocity of each particle. Six variables are required to specify the position and velocity of each particle—three components of the position vector and three components of the velocity vector. Thus, a total of $6N$ microscopic variables is needed to completely specify the mechanical state of an N-particle system. If the values of these $6N$ variables are known at some instant, Newton's laws of motion provide a scheme which, in principle only, allows the mechanical state of the system to be determined at any later time.

Statistical physics is concerned with systems for which N is so large that we are unable, as a practical matter, to specify the mechanical state of the system; that is, there are too many microscopic variables of which to keep track.

Typically, N is of the order of Avogadro's number†

$$N_0 = 6.023 \times 10^{23}$$

The mere contemplation of specifying the mechanical state of such an enormous number of particles ends there—as a very brief contemplation.

Fortunately, complete knowledge of the mechanical state is unnecessary when we are concerned with understanding the macroscopic behavior of the system. The measurable properties reflect the combined actions of many atoms or molecules. The behavior of any particular atom is inconsequential. The gas–piston system of Figure 1.1c is typical of those of interest in statistical physics in that it consists of many particles for which a mechanical description is not feasible and, happily, unnecessary. We resort to either a thermodynamic or a statistical description.

In contrast to the huge number of microscopic variables required to specify a mechanical state, a thermodynamic state is defined by specifying values for only a few macroscopic variables. In general, experiment dictates which are the relevant thermodynamic variables for any specific system. In principle, every measurable property of a system must stand ready to serve as a thermodynamic variable. In practice, the behavior of most systems is adequately described by a small number of variables. Experience also reveals that the thermodynamic variables are not all independent. An equation relating two or more thermodynamic variables is called an *equation of state*. As an example, suppose that the system is a simple gas‡ confined within some definite volume. The thermodynamic states are determined by specifying the density, pressure, volume, and temperature of the gas.§ Only three of these four thermodynamic variables are independent, being related by one equation of state. Equations of state may be established empirically, that is, by fitting an equation to the experimental data. They also may be arrived at theoretically by applying the techniques of kinetic theory or statistical mechanics.

....

1.2.1 The English physicist William Thomson (who later became Lord Kelvin) had a favorite way of impressing upon people the enormity of Avogadro's number. His recipe was this: Imagine a glass of water in which each molecule is "tagged" in such a way as to make it distinguishable. Next, dump this glass of tagged water molecules into the oceans of Earth and stir.

---

† One gram-molecular weight constitutes 1 mole of a substance. According to Avogadro's hypothesis 1 mole of any substance contains the same number of molecules. This number, whose value was unknown to Avogadro, is called *Avogadro's number*.

‡ A simple system is one composed of a single chemical species, as opposed to a mixture of chemically distinct components.

§ More precisely, the equilibrium states are determined by these thermodynamic variables. The notion of thermodynamic equilibrium is discussed in Section 1.3.

Once the tagged molecules have been thoroughly mixed with the untagged water, reach in and dip out a glassful of water. If the mixing has distributed the tagged molecules uniformly, the glass will contain approximately *ten thousand* of the tagged molecules.

Thomson's example is indeed impressive. Devise your own example illustrating the enormity of N_0. The secret of constructing such examples lies in finding some way of diluting N_0 to obtain an easily comprehended number.

We note that by using radioactive isotopes, unknown in Thomson's day, it is possible to tag molecules. The use of these tracer isotopes is now widespread in medicine, biophysics, and ecological studies.

1.2.2 (a) Estimate the number of molecules you inhale in one breath. Your estimate may be rough but indicate the basis for it.

(b) The density of air is approximately 1.3×10^{-3} gm/cm^3. The total mass of our atmosphere is roughly 5×10^{21} gm. Assume that all of the molecules exhaled by George Washington in his dying breath are now distributed uniformly throughout our atmosphere. Make a rough estimate of the number of these molecules which you inhale with each breath.

....

1.3 THERMODYNAMIC EQUILIBRIUM AND TEMPERATURE

Experience has shown that isolated systems tend toward states of *thermodynamic equilibrium*. For a system to be in thermodynamic equilibrium three conditions must be satisfied:

(a) The system must be in mechanical equilibrium, that is, free of any net external force and torque.

(b) The system must be in chemical equilibrium. All chemical reactions must have run their course.

(c) The measurable properties of the system must be spatially uniform throughout and must not change with time.

The uniformity conditions must be amended in situations where the system is intrinsically heterogeneous. For example, a closed insulated container filled with a liquid and its vapor is in thermodynamic equilibrium provided both the vapor and the liquid satisfy the conditions for equilibrium.

....

1.3.1 Give three examples of everyday systems which are in thermodynamic equilibrium or very nearly so. Indicate any features which prevent attainment of complete thermodynamic equilibrium.

....

"Temperature" is a word which means something to every student before he studies thermodynamics. However, the everyday meaning of the word is largely qualitative and is restricted by the range of human sense perceptions. It is necessary to give an operational definition of temperature to render it a useful thermodynamic variable. In the case of temperature it is logically necessary to first give operational status to the idea of temperature *equality*.

Imagine two isolated systems, each in thermodynamic equilibrium. Let thermal contact be established between the two in such a way that the overall combination is isolated. In general, both component systems would undergo changes. When the changes subside the two systems are said to be in *mutual thermodynamic equilibrium*. For simplicity we say the systems are in thermal equilibrium. We also say that their temperatures are equal. Evidently, thermal equilibrium is synonymous with temperature equality.

It should be noted that this definition of temperature equality does not require a scale of temperature. In mechanics Newton's first law of motion presents a similar situation. The first law stands as an operational definition of zero net force. We do not need an operational definition of force (supplied by Newton's second law) to recognize a force-free condition; zero acceleration defines zero force. In a similar fashion, mutual thermodynamic equilibrium defines temperature equality, regardless of what temperature scale is adopted.

1.4 THE ZEROTH LAW OF THERMODYNAMICS

In mathematics the transitive property is introduced as a postulate or as an elementary consequence of other postulates. The transitive law takes the form:

$$\text{If } A = C \text{ and } B = C, \text{ then } A = B.$$

An analogous law holds in thermodynamics, although it is neither a postulate nor a logical truth, but rather the statement of an experimental fact, namely, thermal equilibrium obeys a transitive law. This fact, often referred to as the *zeroth law* of thermodynamics, may be stated as follows: Two systems separately in thermal equilibrium with a third are in thermal equilibrium with each other. Figure 1.2 depicts a schematic demonstration of the zeroth law. Suppose that when thermal contact is established between A and C there are no changes observed in their thermodynamic coordinates[†]; similarly when B and C are placed in thermal contact. If A and B are then placed in thermal contact, it is observed that no changes occur in their thermodynamic states. They are in mutual equilibrium when contact is first established. As Figure 1.2 suggests,

[†] The expressions "thermodynamic coordinate" and "thermodynamic variable" are synonymous.

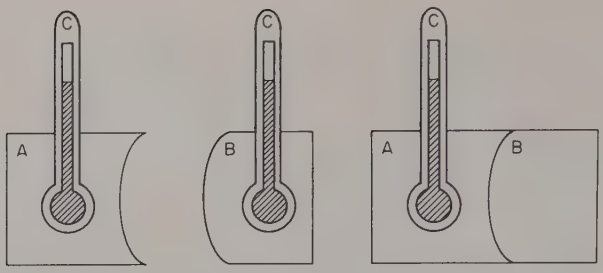


Figure 1.2. Schematic test of the zeroth law of thermodynamics.

C might be a mercury-in-glass thermometer, the height of the mercury column being the thermometric variable. The fact that the height of the column is the same whether the thermometer is in thermal contact with A or B leads one to expect that the thermodynamic state of A and B will remain unchanged when thermal contact is established between them. Experiment bears out the correctness of this expectation thereby establishing the zeroth law of thermodynamics.

Since equality of temperature is synonymous with mutual thermodynamic equilibrium we can phrase the zeroth law in equation form. Even though we have not set about the task of defining a temperature scale we can write

$$T_A = T_C$$

as a symbolic way of indicating that A and C are in thermal equilibrium. This allows the zeroth law to be expressed as the statement:

$$\text{If } T_A = T_C \text{ and } T_B = T_C, \text{ then } T_A = T_B.$$

1.5 PRESSURE—AN IMPORTANT THERMODYNAMIC VARIABLE

Pressure is among the more prominent thermodynamic variables, and as such is encountered throughout thermodynamics and statistical physics. Pressure measures force per unit area. For example, consider an aquarium containing 200 lb of water. If the base of the aquarium has an area of 4 sq ft, the pressure on the base due to the water is

$$P = \frac{200 \text{ lb}}{4 \text{ sq ft}} = 50 \quad \frac{\text{lb}}{\text{sq ft}}$$

The weight of our atmosphere is supported by the surface of the earth. The resulting pressure is approximately 10^6 dyn/cm^2, or roughly 15 lb/sq in. This means that a vertical column of our atmosphere having a cross-sectional area of 1 sq in., weighs about 15 lb. The familiar mercury barometer is generally used to measure the atmospheric pressure. The so-called barometer formula provides the basis for its operation. Consider a column of fluid of height h

and cross-sectional area A. The volume of the fluid is $V = hA$. If ρ denotes the mass density of the fluid, the mass within the column is

$$m = \rho V = \rho hA$$

The weight of the column is

$$W = mg = \rho ghA$$

where g is the acceleration of gravity. The pressure at the base of the column due to the weight of the fluid is

$$\frac{W}{A} = \rho gh \tag{1.1}$$

Notice that A, the area of the column, drops out. Pressure is an intensive variable associated with the extensive variable, force. Equation (1.1) gives the change in pressure produced by the weight of the liquid. If P_0 is the pressure at the surface of the fluid, the total pressure at depth h is

$$P = P_0 + \rho gh \tag{1.2}$$

This is the barometer equation.

....

1.5.1 Estimate the mass and weight of the Earth's atmosphere.

....

The mercury barometer consists of a narrow tube, closed at one end. The tube is filled with mercury and inverted in a reservoir (see Figure 1.3). The

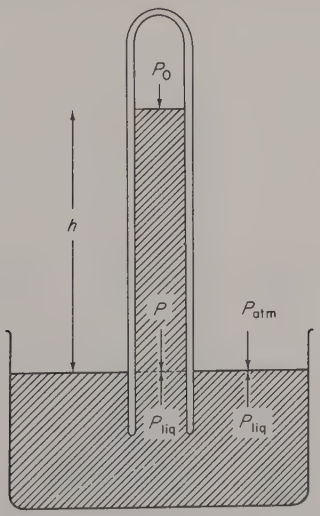


Figure 1.3. Elementary barometer.

pressure at the surface of the mercury reservoir is atmospheric (P_{atm}). Inside the capillary the pressure at the level of the reservoir is $P_0 + \rho gh$. The pressure P_0 at the top of the mercury column is not quite zero because a small amount of mercury vapor fills the otherwise evacuated space. The weight of this vapor produces the pressure P_0, which is ignorable except in precise scientific work. The height of the column is such that the pressures at the level of the reservoir surface are the same inside and outside. This is necessary for mechanical equilibrium. Thus†

$$P_{atm} = P_0 + \rho gh \simeq \rho gh$$

There are several units of pressure currently in use. In addition to the pound per square inch (lb/sq in.) and the dyne per square centimeter (dyn/cm^2), there is the atmosphere (atm)

$$1 \text{ atm} \equiv 1.0132 \times 10^6 \quad \frac{\text{dyn}}{\text{cm}^2}$$

....

1.5.2 At what depth is the total pressure acting on a scuba diver equal to 2 atm?

....

Because pressure can be determined by measuring the height of a barometer column it is often indicated in terms of millimeters of mercury. What is meant by 1 mm of mercury of course is the pressure corresponding to a column of mercury one millimeter in height. The millimeter of mercury gradually is being replaced by the *Torr*. The Torr is an absolute unit of pressure defined as $(1.01325 \times 10^6/760)$ dyn/cm^2, and is very nearly equal to 1 mm of mercury. The bar, defined as 10^6 dyn/cm^2, is another unit of pressure. The millibar (1 mb $= 10^{-3}$ bar) is frequently used in meteorological work. Newspaper weather maps often show isobars (lines of equal pressure) with the pressures indicated in millibars. The pressures indicated are invariably near 1000 mb, that is, near 1 atm.

---

† In addition to the equality sign we shall have occasion to employ the following symbols:
$\equiv$ which is "identically equal to"; (also means "defined as"). For example,

$$x^2 - y^2 \equiv (x-y)(x+y)$$

$$\cosh Z \equiv \frac{e^Z + e^{-Z}}{2}$$

$\simeq$ which is "approximately equal to." For example,

$$\pi \simeq 3.1$$

$\sim$ which is "roughly equal to" or "of the same order of magnitude." For example,

$$\tfrac{4}{3}\pi r^3 \sim r^3$$

.....

1.5.3 A student decides to make a barometer by simply placing the capillary tube in the reservoir, without first filling the tube with mercury. He has heard his instructor explain that the barometer in Figure 1.3 operates because the external (atmospheric) pressure supports the weight of the mercury column. The student argues that if the atmosphere is willing to support the column it should not mind forcing the mercury upward into the column as well. Explain why this will not happen.

1.5.4 The barometer was invented by Torricelli, one of Galileo's students. Torricelli used many different types of liquids, even wine. If the pressure is 1 atm and $g = 980$ cm/sec^2, what is the height of the column if the density of the wine is 0.9 gm/cm^3? This exercise illustrates why one does not see many wine barometers!

.....

Example 1.1 Archimedes' Principle

Archimedes' principle states that an object which displaces a fluid experiences a buoyant force equal to the weight of the fluid it displaces. It is this buoyant force which supports a floating object. To establish the principle, consider a fluid in hydrostatic equilibrium and focus attention on some arbitrary volume element. The net force on the fluid element vanishes when it is in hydrostatic equilibrium. Two types of forces act on the element; a gravitational force equal to the weight of the element and the forces exerted on its surface by the surrounding fluid. The net surface force must just balance the weight of the element in order for it to be in equilibrium. This surface force is a buoyant force—it prevents the fluid element from sinking. If a solid body displaces the fluid element, the same surface forces act on it as acted on the now displaced fluid. Hence, the object experiences a buoyant force equal to the weight of the fluid it displaced.

From Archimedes' principle it follows that an object will float in any liquid whose density is greater than the average density of the object. Ice floats in water because the density of ice (at 0 °C) is 0.92 gm/cm^3. Let ρ_0 denote the average density of an object of volume V which floats in a fluid of density ρ_f. The fraction of the volume V which is submerged is such as to make the buoyant force equal to the weight of the object. We can write

$$mg = \rho_0 Vg = \text{weight of object}$$

If f is the fraction of the volume V submerged, the buoyant force is

$$B = \rho_f fVg$$

Setting $B = mg$ gives

$$f = \frac{\rho_0}{\rho_f}$$

As an example, the density of ice is 0.92 gm/cm^3 and the density of liquid water is 1.00 gm/cm^3. Thus $f = 0.92$, which means that only about 8% of an ice cube or an iceberg is above the surface. ✧

....

1.5.5 The figure of 8% ignores the buoyant force exerted by the air displaced by the exposed part of the ice cube. Explain why this does not introduce any significant error.

1.5.6 (a) Make a rough estimate of your mass density without reference to your volume.

(b) Would you float in a pool of mercury?

(c) Do you think you could walk across a pool of mercury? Consider the question of stability.

1.5.7 The molecular weight of helium is 4 gm/mole. The molecular weight of air is 28.9 gm/mole. A helium-filled balloon rises because it displaces a weight of air in excess of its own weight. Recall that the molar volume is approximately the same for all gases, 22,400 cm^3 at 32 °F and 1 atm (STP).

(a) Estimate the mass and volume of helium required to lift a mass of 100 kg.

(b) In view of your answer for part (a) would it be feasible to market miniblimps? In particular, would a one-man blimp fit in a one-car garage?

1.5.8 A hollow, airtight, brass sphere is suspended from one end of an equal-arm balance. A solid brass cylinder is suspended from the other end, and it is observed that they balance, that is, the balance arm is horizontal. The apparatus is then placed inside a large bell jar and the air is pumped out. Describe and explain what happens to the balance.

....

A pair of entertaining toys rely on Archimedes' principle. Known as the Cartesian diver, the first consists of a hollow plastic figure (the diver) placed in a water-filled bottle. The bottle is fitted with a stopper. Small holes in the side of the diver allow it to be partially filled with water, to the point where it has a density very slightly less than the density of the water and so just floats. This leaves the diver with a small amount of air trapped inside. The bottle is filled so as to leave little or no air space. If the "solid" bottle is squeezed—the

demonstration being more spectacular if only one hand is used—the diver plummets to the bottom. Squeezing the bottle increases the pressure in the water which forces water into the diver, increasing the density of the diver. The diver sinks when his density rises above that of the water. The same results may be achieved by warming the water.

The lava lamp is another device which relies on Archimedes' principle. Large globules of paraffin are submerged in a cylinder filled with machine oil. The density of the paraffin at room temperature is just slightly greater than that of the oil at the same temperature. An electric light bulb illuminates and heats the base of the cylinder. The temperature of the oil is greatest at the bottom of the cylinder. The heat causes the paraffin to expand more than the oil, to the point where it becomes less dense than the oil, whereupon it rises toward the surface. As it rises it is cooled by the surrounding oil. It contracts until it becomes more dense than the surrounding oil, whereupon it sinks. The cohesive nature of the paraffin results in a strking slow motion rise and fall.

1.6 TEMPERATURE AND THERMOMETRY

We wish to move the notion of temperature from its status as a sense perception to one of a reliable and useful scientific concept. This requires an operational definition of temperature, that is, we must specify how temperature is measured.

Thermometers are characterized by a thermometric substance and a thermometric variable. For example, in the familiar household thermometer, the thermometric substance is a liquid, usually colored alcohol or mercury. The length of the liquid column is the thermometric variable.

The most fundamental scale of temperature is based on the laws of thermodynamics. A discussion of this scale, known as the thermodynamic scale, is presented in Chapter 5. The upshot of the development presented there is this: On the thermodynamic scale, temperature is defined in terms of heat. A temperature so defined is independent of the properties peculiar to any thermometric substance. This is clearly a most desirable characteristic. The recorded temperature depends on the state of the system but not on the material used as a thermometric substance.

An empirical scale is arrived at by arbitrarily defining a temperature in terms of some thermometric variable. A mathematically elementary relationship is generally adopted. For the constant volume gas thermometer (Figure 1.4) and several others, the relation is linear. The temperature is defined in terms of the gas pressure P by

$$T(P) = bP$$

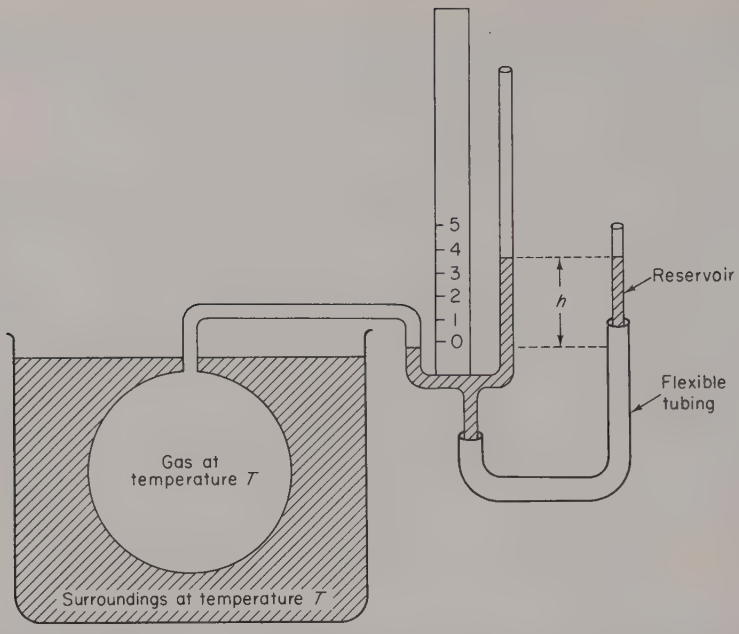


Figure 1.4. Schematic of a constant volume gas thermometer. The gas is maintained at a constant volume by means of the mercury reservoir. If the temperature of the bath T rises, the level of the reservoir h must be increased to maintain the same volume. In practice many refinements and corrections are necessary. Among these are accounting for the change in volume of the gas container and the fact that the gas in the tube leading to the manometer is not at the same temperature as the gas in the main vessel.

More generally we can write

$$T(X) = bX \tag{1.3}$$

as a linear relation defining a temperature $T(X)$ in terms of the value of some thermometric variable X. The value of the proportionality constant b is fixed by assigning a temperature to some standard fixed point.

By virtue of international agreement, the temperature of water at its triple point[†] has been assigned the value of 273.16 K, where K is the abbreviation for the unit called the kelvin. Accessibility and ease of reproduction are the primary virtues of using the triple point of water as the fundamental fixed point.

[†] The solid, liquid, and gaseous phases of water can coexist in thermodynamic equilibrium at a unique pressure (4.58 Torr) and temperature. This is referred to as the triple point of water.

If X_3 denotes the value of the thermometric variable at the triple point of water, we have

$$273.16\,\mathrm{K} = bX_3$$

or

$$b = \frac{273.16\,\mathrm{K}}{X_3}$$

The empirical temperature scale is now defined as

$$T(X) = 273.16\,\mathrm{K}\,\frac{X}{X_3} \tag{1.4}$$

In particular, for the constant volume gas thermometer,

$$T(P) = 273.16\,\mathrm{K}\,\frac{P}{P_3} \tag{1.4a}$$

....

1.6.1 Using a constant volume gas thermometer of modest precision, a student finds a value of 1.37 for the ratio of the steam-point pressure and the triple-point pressure. What is the corresponding steam-point temperature?

....

The constant volume gas thermometer is of special interest because it can be used to establish the fundamental thermodynamic scale over an important range of temperatures. Constant volume gas thermometers, employing different gases, generally will not be in agreement except at the triple point of water. Further, at any temperature other than that of the triple point of water, the thermometer reading depends slightly on the gas pressure at the triple point P_3. For example, suppose that oxygen is used as the gas, and that the thermometer is filled so as to give a triple-point pressure of 200 Torr ($P_3 =$ 200 Torr). If the thermometer is then used to determine the steam point[†] via (1.4a),

$$T_{\mathrm{stm}} = 273.16\,\mathrm{K}\,\frac{P_{\mathrm{stm}}}{P_3}$$

a value near, but in excess of, 373.15 K would be recorded. Suppose next that some of the oxygen is removed until the triple-point pressure is 100 Torr. The

[†] The boiling point is the temperature at which a liquid and its vapor are in thermodynamic equilibrium. The normal boiling point is the equilibrium temperature when the pressure of the vapor is exactly 1 atm. The normal boiling point of pure water is called the steam point.

temperature of the steam point, computed according to the above ralation, would be different than the value found with $P_3 = 200$ Torr. It would, in fact, be nearer 373.15 K. If the triple-point pressure is reduced again and the steam point determined, the temperature recorded is again different, and still nearer 373.15 K. A sequence of such measurements, using progressively lower triple-point pressures enables one to extrapolate the results to the point where $P_3 = 0$. It is, of course, impossible to operate a constant volume gas thermometer at zero pressure. The extrapolated value of T_{stm}, using oxygen as the thermometric substance, is 373.15 K. If a different gas, say helium or nitrogen, is used instead of oxygen, a similar variation in T_{stm} is noted as the triple-point pressure is reduced. However, and this is the vitally significant point, when the results for helium, nitrogen, and so on, are extrapolated to $P_3 = 0$, they give a common temperature for the steam point, namely 373.15 K. Extrapolating the data to the point where $P_3 = 0$ gives a value for T_{stm} which is the same regardless of the gas used. This is precisely the requirement for a thermodynamic scale—a thermometer whose readings are independent of the thermometric substance. Figure 1.5 suggests the relationship between the values of P_3 and T_{stm}, as given by

$$T_{stm} = 273.16\,\text{K}\,\frac{P_{stm}}{P_3}$$

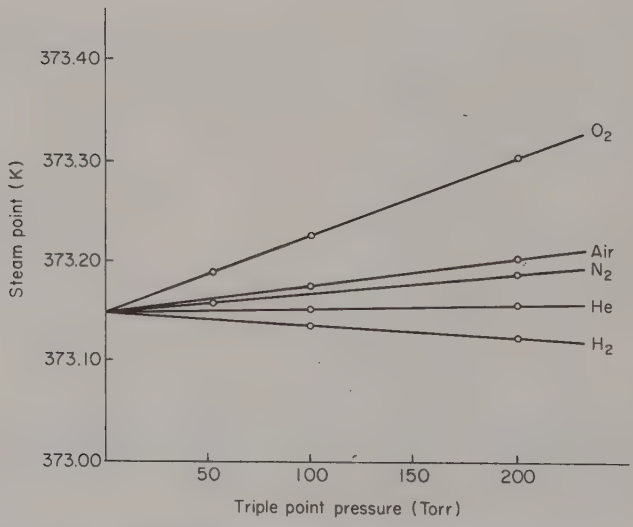


Figure 1.5. The steam-point temperature recorded by a constant volume gas thermometer. For any particular gas, the steam-point temperature depends on the triple-point pressure. However, when the data are extrapolated to zero pressure a unique steam point 373.15 K is indicated. [Adapted from D. Halliday and R. Resnick, "Physics." Wiley, New York (1960).]

A common temperature for the steam point is reached only when the data are extrapolated to $P_3 = 0$.

The temperature resulting from such extrapolation is often referred to as the ideal gas temperature

$$T_{\text{I.G.}} = \lim_{P_3 \to 0} 273.16\,\text{K}\,\frac{P}{P_3} \tag{1.5}$$

where I.G. stands for ideal gas. The concept of an ideal gas stems from the fact that all real gases exhibit similar behaviors at sufficiently low pressures. The ideal gas equation of state has the form

$$PV = nRT \tag{1.6}$$

where P, V, and T are the gas pressure, volume, and temperature, respectively and n denotes the number of moles of gas and R is the universal gas constant. In the cgs system of units R has the value

$$R = 8.32 \times 10^7\ \text{erg/K} \tag{1.7}$$

An alternate form of the ideal gas equation of state employs the number of molecules N rather than the number of moles. In (1.6) we replace n by the ratio of the total number of molecules and Avogadro's number

$$n = \frac{N}{N_0} \tag{1.8}$$

If, further, we define a constant k by

$$k = \frac{R}{N_0} \tag{1.9}$$

the ideal gas equation of state becomes

$$PV = NkT \tag{1.10}$$

The quantity k is called the Boltzmann constant and has the value

$$k = 1.38 \times 10^{-16}\ \text{erg/K}$$

....

1.6.2 Let M and ρ denote the molecular weight and mass density of a gas, respectively. Show that the ideal gas equation of state can be written as

$$P = \frac{R\rho T}{M}$$

....

In addition to the Kelvin scale, three others are in widespread use today. These are the Celsius, Fahrenheit, and Rankine scales. On the Celsius scale, temperature t_C is related to the Kelvin temperature T by

$$t_C = (T - 273.15)\,°C \tag{1.11}$$

with °C being the abbreviation for degrees Celsius. Thus, for example, the Celsius temperature of the triple point of water is 0.01 °C, since $T_3 = 273.16\,K$. Notice that the kelvin and the Celsius degree have the same size. The scales differ only in the location of their zeros.

Throughout most English-speaking countries, the Fahrenheit scale of temperature is widely used. The symbol °F on the household thermometer stands for degrees Fahrenheit.

Temperatures on the Fahrenheit scale are related to those of the Celsius scale by the equation

$$t_F = \left(32 + \frac{9}{5} t_C\right)°F \tag{1.12}$$

This places the ice and steam points at 32 °F and 212 °F. The Fahrenheit and Celsius scales evidently differ in the size of their units and in the location of their zeros.

....

1.6.3 Derive the inverse of (1.12), that is, the equation expressing t_C in terms of t_F.

....

The Rankine temperature scale is widely used in engineering. The Rankine and Kelvin scales have a common zero (absolute zero), but the size of their units differ. The Rankine degree is equal to the Fahrenheit degree, and thus is only 5/9 as large as the kelvin and Celsius degree,

$$1\,°R = \frac{5}{9} K \tag{1.13}$$

As a consequence, a temperature on the Rankine scale is 9/5 as great as the corresponding Kelvin temperature

$$t_R = \left(\frac{9}{5} T\right)°R \tag{1.14}$$

For example, the Rankine temperature of the ice point of water is

$$\frac{9}{5}(273.15)\,°R = 491.57\,°R$$

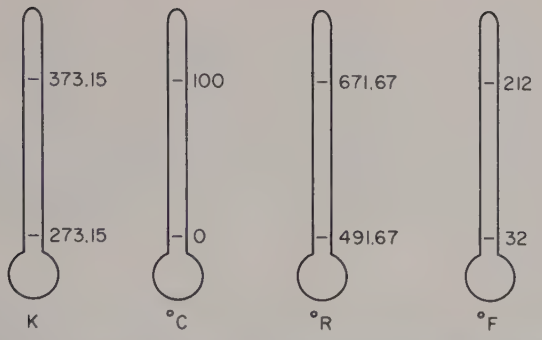


Figure 1.6. Four identically constructed thermometers calibrated in kelvins (K), degrees Celsius (°C), Rankine (°R), and Fahrenheit (°F). The figures indicate the temperature values for the ice and steam points.

The Rankine scale bears the same relationship to the Fahrenheit scale as does the Kelvin scale to the Celsius scale, that is, their temperature units are of equal size but the locations of the origins differ. Figure 1.6 summarizes the relationships among the four scales.

....

1.6.4 Determine the Fahrenheit temperature of absolute zero (0 °R).

1.6.5 Derive formulas expressing

(a) Rankine temperature in terms of Celsius temperature, and
(b) Fahrenheit temperature in terms of Kelvin temperature.

....

REFERENCES

Three splendid texts which have been favorites for many years are:

E. Fermi, "Thermodynamics." Dover, New York (1956).

M. W. Zemansky, "Heat and Thermodynamics," 5th ed. McGraw-Hill, New York (1968).

This edition includes material on kinetic theory and statistical mechanics not found in earlier ones.

G. N. Lewis and M. Randall, "Thermodynamics," 2nd ed. McGraw-Hill, New York (1961).

Strongly recommended supplements, aimed at the novice, are:

H. C. Van Ness, "Understanding Thermodynamics." McGraw-Hill, New York (1969).

and

M. W. Zemansky, "Temperatures Very Low and Very High." Van Nostrand-Reinhold, Princeton, New Jersey (1964).

Brief glimpses of the underlying atomic structure of matter and its relation to the macroscopic world of our senses are presented in:

M. Born, "The Restless Universe," Chapter 1. Dover, New York (1951).

and

R. P. Feynman, R. B. Leighton, and M. Sands, "The Feynman Lectures on Physics," Volume 1, Chapters 1 and 2. Addison-Wesley, Reading, Massachusetts (1963).

A problem-solver volume is:

L. Pincherle, "Worked Problems in Heat, Thermodynamics, and Kinetic Theory for Physics Students." Pergamon, Oxford (1966).

An authoritative and exhaustive trio covering virtually every aspect of thermometry is:

"Temperature, Its Measurement and Control in Science and Industry," Volume III. Van Nostrand-Reinhold, Princeton, New Jersey (1962).
F. G. Brickwedde (ed.), Part 1. Provides detailed expositions of thermometry scales and methods.
A. I. Dahl (ed.), Part 2. Is devoted to thermometry principles and instrumentation.
J. D. Hardy (ed.), Part 3. Deals with applications in the fields of biology and medicine.

Many practical aspects of thermometry, including the most recent international standards of temperature, are reviewed in:

W. T. Gray and D. I. Finch, How Accurately Can Temperature Be Measured? *Phys. Today* **24**, 32 (Sept. 1971).

Chapter 2

Mathematical Preliminaries

This chapter is to serve as an introduction or as a review of the mathematics needed for the development of thermodynamics. Sections 2.1 and 2.2 present the concepts of the differential and the partial derivative, and stress their geometrical significance. Two commonly-used partial derivative relations are derived and illustrated in Section 2.3. In Section 2.4 the coefficient of expansion and the compressibility are introduced to illustrate one way in which partial derivatives are used in thermodynamics.

2.1 DIFFERENTIALS

On a macroscopic scale the observable properties of systems are generally continuous. For example, the pressure produced by the atmosphere of the Earth at a given locality varies smoothly with altitude. The temperature of the atmosphere also exhibits spatial continuity. The continuity of physical properties allows a very concise mathematical formulation of thermodynamics. Two mathematical constructs are especially useful: the differential and the partial derivative. Our aim here is to introduce these two quantities and establish their geometric significance. An appreciation of their geometrical content greatly strengthens their physical interpretations.

The differential is useful in relating changes of two or more interdependent

quantities. Let x and $y(x)$ denote corresponding values of the independent and dependent variables, respectively. We write

$$\Delta y = y(x+dx) - y(x) \tag{2.1}$$

for the change in y resulting from a change of dx in x. We call dx the differential of x. It denotes a change in the value of the variable x. The size of dx is arbitrary—it need not be small in any sense of the word. The differential of the dependent variable is defined by

$$dy \equiv \left(\frac{dy}{dx}\right) dx \tag{2.2}$$

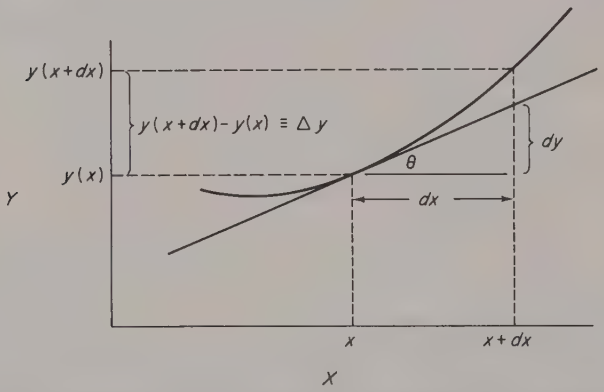


Figure 2.1. The differential dy is a linear approximation to the change Δy, with $dy = \tan\theta\, dx = (dy/dx)\, dx$.

Parentheses have been used to indicate that (dy/dx) is the derivative of y with respect to x. Reference to Figure 2.1 reveals the geometrical significance of dy. The differential dy gives a linear approximation to the change Δy. For sufficiently small dx, dy is a suitable replacement for Δy

$$\Delta y \simeq \left(\frac{dy}{dx}\right) dx \tag{2.3}$$

What is the criterion for "sufficiently small"? An example will help to establish a feeling for this.

Example 2.1 The Fractional Change in g

Treat the Earth as a sphere of radius R, having a uniform distribution of mass. The gravitational acceleration a distance r from the center of the Earth, and above its surface, is

$$g(r) = g_0 \frac{R^2}{r^2}$$

where g_0 is the acceleration of gravity at the surface. We have

$$\left(\frac{dg}{dr}\right) = -2g_0\frac{R^2}{r^3} = \frac{-2}{r}g(r)$$

so that the differential of g is

$$dg = -2g(r)\frac{dr}{r}$$

Observe that $dg/g = -2\,dr/r$. If the fractional change in r, dr/r, is much less than unity, so will dg/g be. For small dr/r then, we expect that $\Delta g/g$, the fractional change in g, is adequately approximated by dg/g. A direct calculation of $\Delta g/g$ bears out this expectation.

$$\Delta g = g(r+dr) - g(r) = g_0 R^2\left\{\frac{1}{(r+dr)^2} - \frac{1}{r^2}\right\}$$

Minor algebra produces

$$\frac{\Delta g}{g} = -2\frac{dr}{r}\frac{\left(1+\frac{1}{2}\frac{dr}{r}\right)}{\left(1+\frac{dr}{r}\right)^2} = \frac{dg}{g}\frac{\left(1+\frac{1}{2}\frac{dr}{r}\right)}{\left(1+\frac{dr}{r}\right)^2}$$

As expected $\Delta g/g$ is closely approximated by dg/g when dr/r is much less than unity. Because of the continuous nature of $g(r)$ we can conceive of arbitrarily small values of dr, whereupon dg represents the change in g. ✧

The example illustrates a general relationship between differentials and differences, namely that the change in y for sufficiently small dx may be represented by the differential

$$dy = \left(\frac{dy}{dx}\right)dx \simeq \Delta y \tag{2.4}$$

Thermodynamics is concerned with situations where x and $y(x)$ represent measurable quantities, that is, physical properties of matter. The continuous nature of such properties allows us to envision changes sufficiently small to permit the use of

$$dy \equiv \left(\frac{dy}{dx}\right)dx \tag{2.2}$$

to relate changes in x and y.

.....

2.1.1 Determine the relation between $\Delta y/y$ and dy/y for the parabola $y = x^2$. If dx/x is 0.01, by what amount do $\Delta y/y$ and dy/y differ?

.....

2.2 PARTIAL DERIVATIVES

The need for partial derivatives arises when more than one independent variable enters. Consider a function z of two independent variables, x and y,

$$z = z(x, y)$$

Geometrically $z(x, y)$ may be thought of as the height of a surface above the x–y plane. A three-dimensional surface can be represented by an equation of the form $f(x, y, z) = 0$. For example, the equation of a sphere of radius b centered at the origin may be written as

$$f(x, y, z) \equiv x^2 + y^2 + z^2 - b^2 = 0$$

Here we regard z as the dependent variable. However, we are free to choose any pair as independent and the third as dependent.

We can form derivatives of $z(x, y)$ with respect to both x and y. The derivatives are called *partial* derivatives and are defined by

$$\left(\frac{\partial z}{\partial x}\right) \equiv \lim_{\Delta x \to 0} \frac{z(x+\Delta x, y) - z(x, y)}{\Delta x} \tag{2.5}$$

$$\left(\frac{\partial z}{\partial y}\right) \equiv \lim_{\Delta y \to 0} \frac{z(x, y+\Delta y) - z(x, y)}{\Delta y} \tag{2.6}$$

In forming $(\partial z/\partial x)$, the definition (2.5) shows that the independent variable y is treated as a constant. As an example, consider the function

$$z(x, y) = 2x^2 y$$

where x and y are independent variables. The partial derivatives of $z(x, y)$ are

$$\frac{\partial z}{\partial x} = 4xy, \qquad \frac{\partial z}{\partial y} = 2x^2$$

.....

2.2.1 Verify for the function

$$z(x, y) = 2x^2 y$$

that

$$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)$$

2.2.2 For the functions in (i) and (ii) below

(a) evaluate $\partial f/\partial x$ and $\partial f/\partial y$,
(b) verify that

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

that is, show that the order of partial differentiation does not matter.

(i) $f(x,y) = x^2y - xy^2$,
(ii) $f(x,y) = ye^x$.

....

To avoid confusion the partial derivatives are often subscripted to indicate which variables are independent of the differentiation variable. In the previous example we might have written

$$\left(\frac{\partial z}{\partial x}\right)_y = 4xy, \qquad \left(\frac{\partial z}{\partial y}\right)_x = 2x^2$$

The subscripts tell us which variables are being treated as constant during the differentiation. Figure 2.2 suggests the geometrical significance of the partial derivative.

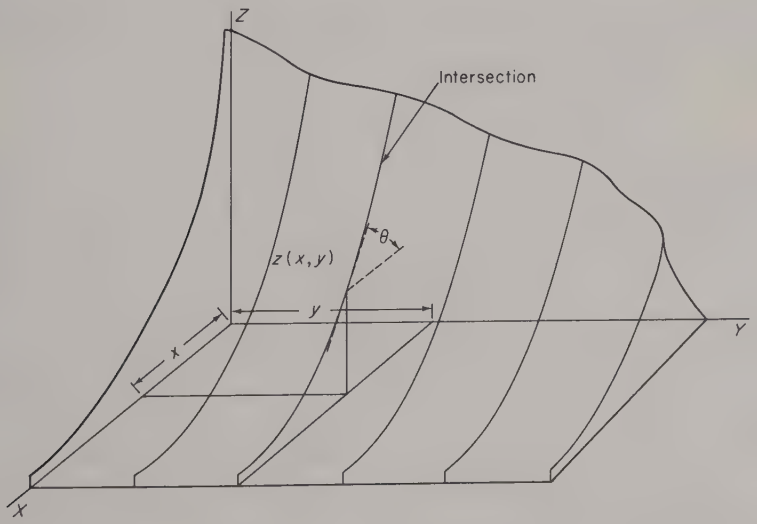


Figure 2.2. The partial derivative $(\partial z/\partial x)_y = \tan\theta$ gives the slope of a line formed by the intersection of the plane $z = z(x,y)$ and the plane $y =$ constant.

....

2.2.3 With reference to Figure 2.2, explain the geometrical significance of $(\partial z/\partial y)_x$.

....

We define the differential of z by

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy \tag{2.7}$$

It follows from the geometrical significance of the partial derivatives that, for sufficiently small dx and dy, dz gives the change in z between $z(x, y)$ and $z(x+dx, y+dy)$. Thus

$$\Delta z \equiv z(x+dx, y+dy) - z(x, y) \simeq dz \tag{2.8}$$

To demonstrate this more convincingly we subtract and add $z(x, y+dy)$ on the right side, forming two differences,

$$\Delta z = \{z(x+dx, y+dy) - z(x, y+dy)\} + \{z(x, y+dy) - z(x, y)\}$$

For sufficiently small dx and dy the differences may be replaced by the corresponding differentials. Thus

$$\Delta z \simeq \left(\frac{\partial z}{\partial x}\right)_{y+dy} dx + \left(\frac{\partial z}{\partial y}\right)_x dy$$

The first derivative is evaluated at $y+dy$. For small dy this differs from the derivative at y by an amount proportional to dy. Without sacrificing the accuracy of the approximation we can set

$$\left(\frac{\partial z}{\partial x}\right)_{y+dy} dx \simeq \left(\frac{\partial z}{\partial x}\right)_y dx$$

which establishes the desired result

$$\Delta z \simeq \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy = dz \tag{2.8a}$$

As an illustration, consider

$$z(x, y) = yx^3$$

We have

$$\left(\frac{\partial z}{\partial x}\right)_y = 3x^2 y, \qquad \left(\frac{\partial z}{\partial y}\right)_x = x^3$$

whereupon (2.7) gives

$$dz = 3x^2 y\, dx + x^3\, dy$$

.....

2.2.4 The entropy S of an ideal gas is given in terms of the pressure and temperature by†

$$S(T,P) = C_p \ln T - R \ln P + S_0$$

Show that

$$dS = \frac{C_p}{T} dT - \frac{R}{P} dP$$

.....

2.3 TWO IMPORTANT PARTIAL DERIVATIVE RELATIONS

Let $z(x,y)$ denote a function of the independent variables x and y. We can write

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy \tag{2.9}$$

We are equally free to regard x as a function dependent on the independent variables z and y. The differential of $x(z,y)$ is then expressible as

$$dx = \left(\frac{\partial x}{\partial z}\right)_y dz + \left(\frac{\partial x}{\partial y}\right)_z dy \tag{2.10}$$

Both (2.9) and (2.10) hold simultaneously, that is, both are simply alternate ways of relating changes in the interdependent trio x, y, and z. If (2.10) is inserted into (2.9), there results

$$dz = \left\{\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y\right\} dz + \left\{\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z + \left(\frac{\partial z}{\partial y}\right)_x\right\} dy \tag{2.11}$$

This relation is valid for all conceivable changes in y and z. If y is held constant while z varies, so that $dy = 0$, $dz \neq 0$, (2.11) yields the first of the two sought after relations

$$\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y = 1 \tag{2.12}$$

An alternate form

$$\left(\frac{\partial z}{\partial x}\right)_y = \frac{1}{\left(\frac{\partial x}{\partial z}\right)_y} \tag{2.13}$$

---

† We denote by $\ln x$ the logarithm of x to the *natural base e*. Any other base will be specifically noted, for example, $\log_{10} x$, $\log_2 x$, and so on.

reveals it as an inversion theorem. For a situation where z remains constant, but for which y changes, $dz = 0$ and $dy \neq 0$, and (2.11) becomes

$$0 = \left\{ \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z + \left(\frac{\partial z}{\partial y}\right)_x \right\} dy$$

Since dy is not zero, the expression in the braces must vanish. This gives the second relation, a pseudo chain rule,†

$$\left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z = -\left(\frac{\partial z}{\partial y}\right)_x \tag{2.14}$$

An alternate form, obtained by applying (2.12), is

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \tag{2.15}$$

Note that each of the derivatives in (2.15) involves the three variables in "cyclic order"

....

2.3.1 Convert (2.14) into (2.15) by making use of (2.12).

2.3.2 Verify the relation

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1$$

for the variables x, y, and z related by

$$x^2 + y^2 + z^2 - a^2 = 0$$

Hint: Show that $(\partial x/\partial y)_z = -y/x$ and then evaluate the other two partial derivatives either by inspection or by a cyclic permutation argument.

....

† Recall that if $g = g(x)$ and $x = x(y)$, then the usual chain rule of differential calculus gives

$$\left(\frac{dg}{dy}\right) = \left(\frac{dg}{dx}\right)\left(\frac{dx}{dy}\right)$$

2.4 USING PARTIAL DERIVATIVES IN THERMODYNAMICS

Evaluating partial derivatives of a specified mathematical function is a routine operation. Manipulating partial derivatives in calculations is a skill polished by practice. In applications of thermodynamics it often is necessary to use quantities which are represented by partial derivatives. The coefficient of volume expansion is one such quantity. Most substances undergo a change in volume when they are heated at constant pressure. The volume of a simple fluid or solid may be regarded as a variable dependent on the two independent variables pressure and temperature

$$V = V(P, T)$$

If the temperature is changed slightly, from T to $T+\Delta T$, while the pressure remains constant, the volume of the sample changes by an amount proportional to ΔT. The change in volume is also proportional to the volume itself. We can therefore express the change in volume as

$$\Delta V = \beta V \Delta T \tag{2.16}$$

The proportionality factor β is termed the coefficient of volume expansion. From the experimental viewpoint, β is given by

$$\beta = \frac{1}{V}\frac{\Delta V}{\Delta T} \tag{2.17}$$

The measured ratio $\Delta V/\Delta T$ corresponds to the partial derivative $(\partial V/\partial T)_P$. More precisely,

$$\left(\frac{\partial V}{\partial T}\right)_P = \lim_{\Delta T \to 0} \frac{\Delta V}{\Delta T}$$

For small changes in T and V then, the partial derivative $(\partial V/\partial T)_P$ is adequately represented by the measured ratio $\Delta V/\Delta T$, and vice versa. Consequently, in theoretical developments the coefficient of volume expansion is identified as the quantity

$$\beta = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P \tag{2.18}$$

A quantity closely related to β is the coefficient of linear expansion. If $L(T, F)$ is the length of an object subject to a force F, the coefficient of linear expansion is defined as

$$\alpha = \frac{1}{L}\left(\frac{\partial L}{\partial T}\right)_F \tag{2.19}$$

Comparing this with (2.18) we see that α is the linear analog of β; length appears in place of the volume, force in place of pressure. Experimentally, α is determined by measuring the temperature variations in length. If the length of the sample changes from L to $L+\Delta L$ as the temperature changes from T to $T+\Delta T$,

$$\alpha = \frac{1}{L}\frac{\Delta L}{\Delta T} \tag{2.20}$$

In practice, (2.20) is used to relate the changes in length and temperature, that is,

$$\Delta L = \alpha L \Delta T \tag{2.21}$$

....

2.4.1 In 1956, a high-school athlete hurled a discus 185 ft $2\frac{3}{8}$ in. at the Ohio State Championship meet. The distance was measured with a tape made of invar, a high-grade steel, with the (very small) coefficient of expansion $\alpha = 0.7 \times 10^{-6}/°C$. At that time the National High School record was 185 ft $2\frac{1}{2}$ in. The tape was calibrated at a temperature of 25 °C. The temperature on the day of the track meet was 85 °F. Because of thermal expansion the actual distance was greater than that measured with the tape. Determine if the actual distance exceeded the national record.

2.4.2 A brass sphere ($\alpha = 1.9 \times 10^{-5}/°C$) has a diameter of 5.00 cm at 25 °C. A steel ring ($\alpha = 1.1 \times 10^{-5}/°C$) has a diameter of 4.995 cm at 25 °C. At what temperature will the sphere just be able to pass through the ring

(a) if the ring is heated, but not the sphere?
(b) if both the ring and the sphere have the same temperature?

2.4.3 Derive the result $\beta = 3\alpha$ by considering the change in volume of a cube of side L.
Hint: A quick derivation makes use of

$$V = L^3, \qquad dV = 3L^2\, dL$$

2.4.4 Determine the change in density of a solid copper sphere when its temperature rises from 20 °C to 60 °C. The density of copper at 20 °C is 8.96 gm/cm^3. You may wish to derive first the result

$$\frac{\Delta\rho}{\rho} = -\beta \Delta T$$

where $\Delta\rho$ and ΔT are the changes in density and temperature.

....

TABLE 2.1

Coefficients of Linear Expansion for Selected Liquids and Solids[a]

Liquids		Solids	
Substance	$10^3\alpha$ (°C^{-1})	Substance	$10^5\alpha$ (°C^{-1})
Acetic acid	1.07	Silver (Ag)	1.89
Acetone	1.49	Aluminum (Al)	2.80
Benzene	1.24	Gold (Au)	1.42
Carbon		Calcium (Ca)	2.50
tetrachloride	1.24	Iron (Fe)	1.17
Ether	1.66	Sodium (Na)	7.1
Mercury	0.18	Lead (Pb)	2.91
Pentane	1.61	Rubidium (Rb)	9.0
Water	0.21	Zinc (Zn)	3.3
		Carbon steel	1.15
		Glass	0.90

[a] Values refer to measurements at a pressure of 1 atm and for temperatures near 20 °C.

$$\alpha \equiv \frac{1}{L}\left(\frac{\partial L}{\partial T}\right)_F$$

Values for liquids adapted from J. H. Perry, "Chemical Engineer's Handbook." Copyright by McGraw-Hill, New York (1950). Values for solids taken from "International Critical Tables of Numerical Data: Physics, Chemistry, and Technology." Nat. Res. Council, Washington, D.C., 1923–1933.

For solids, measurements of α are much simpler than determinations of β. Because of this, and in view of the simple proportionality of β and α,[†] values of β for solids are seldom tabulated. Typical values of α are listed in Table 2.1. It will be noted that the listed values of α are positive and that they refer to a common temperature of 20 °C. For most solids α varies smoothly with temperature and is nearly independent of pressure changes. The tabulated values of α cluster near 10^{-5}/°C for solids and 10^{-3}/°C for liquids. This means that the volume of a solid changes by about one part in 10^5 for each degree change in temperature. A liquid expands by about one part in 10^3 for the same change in temperature. A few substances exhibit negative values of α and β over certain ranges of temperature. Most notable of these is ice for which β is negative at temperatures below 63 K and (for water) over the narrow range of

[†] For isotropic solids $\beta = 3\alpha$ (see Problem 2.4.3).

temperature from 0 °C to 3.98 °C. Thus, if water at 1 °C is warmed to 2 °C, its volume decreases and its density increases. The maximum density of water occurs at a temperature of 3.98 °C (at a pressure of 1 atm). Type metal, an alloy of lead and antimony, also exhibits a negative value of β. After the type is formed the mold is inverted and heated slightly. The type shrinks and falls free of the mold. At one time a similar alloy was used for plugs in ceiling fire sprinkler systems.

....

2.4.5 Explain why ice forms first at the surface of a lake rather than at its bottom.

2.4.6 Suppose that ice did not float and reflect on some possible consequences.

....

For gases, β is strongly temperature-dependent, being very nearly equal to the reciprocal of the Kelvin temperature

$$\beta_{\text{gas}} \simeq \frac{1}{T}$$

For the ideal gas, β is exactly $1/T$ (compare Problem 2.4.8).

Another quantity related to a partial derivative is the isothermal compressibility κ. The analytical definition of κ is

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \tag{2.22}$$

Because $(\partial V/\partial P)_T$ is negative for most substances (they shrink when squeezed) the minus sign is inserted to make κ a positive quantity. Thus κ measures the extent to which the volume of a system is diminished by an increase in pressure, hence the name compressibility. Experimentally, κ is determined as the ratio

$$\kappa = -\frac{1}{V}\frac{\Delta V}{\Delta P} \tag{2.23}$$

In practice, (2.23) is used to relate the changes in volume and pressure

$$\Delta V = -\kappa V \Delta P \tag{2.23a}$$

....

2.4.7 A student of mass 100 kg balances himself on a cube of lead (Pb). The volume of the cube is 1 cm^3 before the students stands on it. What is the fractional change in its volume while he balances atop it?

....

TABLE 2.2

Compressibilities of Selected Liquids and Solids[a]

Liquids		Solids	
Substance	$10^{12}\kappa$ (cm^2/dyn)	Substance	$10^6\kappa$ (atm^{-1})
Acetic acid	0.91	Aluminum (Al)	1.38
Acetone	1.28	Aluminum oxide (Al_2O_3)	0.40
Benzene	0.94	Gold (Au)	0.60
Carbon disulfide	0.93	Calcium (Ca)	5.89
Carbon tetrachloride	1.05	Calcium carbonate ($CaCO_3$)	1.36
Ether	1.87	Flourite (CaF_2)	1.23
Ethyl bromide	1.29	Iron (Fe)	0.61
Mercury	0.41	Pyrite (FeS_2)	1.70
Methanol	1.21	Lithium (Li)	8.98
Toluene	0.90	Lithium fluoride (LiF)	1.53
Water	0.46	Lead (Pb)	2.38
		Galena (PbS)	1.91
		Rubidium (Rb)	40.5
		Zinc (Zn)	1.76

[a] Values refer to pressures near 1 atm. The reference temperature is 20°C for liquids and 30°C for solids. Since 1 atm $= 1.013 \times 10^6$ dyn/cm^2, 10^{-6} atm^{-1} is very nearly equal 10^{-12} cm^2/dyn and the numerical values for liquids and solids are directly comparable

$$\kappa \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$$

Values for liquids taken from "Handbook of Chemistry and Physics," 51st ed. Chem. Rubber Publ. Co., Cleveland, Ohio (1970). Values for solids taken from "International Critical Tables of Numerical Data: Physics, Chemistry, and Technology." Nat. Res. Council, Washington, D.C., 1923–1933.

Typical values of κ are displayed in Table 2.2. It is evident from the table that 10^{-12} cm^2/dyn $\simeq 10^{-6}$/atm is a typical value of κ. A pressure increase of 1 atm therefore produces a fractional change in volume of approximately one part in a million. For gases the compressibility is very nearly the reciprocal of the pressure

$$\kappa_{\text{gas}} \simeq \frac{1}{P}$$

In fact, for the ideal gas κ is exactly $1/P$ (compare Problem 2.4.8). For a pressure of 1 atm we have

$$\kappa_{\text{gas}} \simeq 1 \text{ atm}^{-1}$$

about one million times greater than the compressibility of a typical solid. The relatively large compressibilities of gases is in accord with everyday experience. A balloon filled with air or helium is readily compressed; a steel ball bearing seems virtually incompressible. The fact that κ_{gas} is inversely proportional to the pressure is also consistent with experience. A half-flat tire is deformed much more by the weight of an auto than one which is properly inflated.

....

2.4.8 Use the ideal gas equation of state $PV = nRT$ to show that

$$\beta_{\text{I.G.}} = \frac{1}{T}, \qquad \kappa_{\text{I.G.}} = \frac{1}{P}$$

....

Another derivative of physical significance is $(\partial P/\partial T)_V$, which measures the change in pressure produced by a change in temperature at constant volume. It would be extremely difficult to measure $(\partial P/\partial T)_V$ directly. Fortunately, it is also unnecessary. By applying the pseudo chain rule (2.14) and the inversion theorem (2.13) one finds

$$\left(\frac{\partial P}{\partial T}\right)_V = \frac{\beta}{\kappa} \tag{2.24}$$

This is typical of the way the pseudo chain rule and the inversion theorem are used in thermodynamics. A quantity difficult to measure is expressed in terms of more readily observable parameters.

REFERENCES

A concise introduction to much of the required mathematics is provided in:

M. H. Hull, Jr., "The Calculus of Physics." Benjamin, New York (1969).

The derivative concept is covered in Chapter 1. Partial differentiation receives attention in Chapter 5.

An abbreviated "self-propelled" volume which offers a concise introduction or rapid review of calculus fundamentals is:

D. Kleppner and N. Ramsey, "Quick Calculus." Wiley, New York (1965).

Chapter 3

The First Law of Thermodynamics

3.1 THE MANY FACES OF ENERGY

Energy is a concept which shows us many faces. Mechanics has acquainted us with the kinetic energy of a moving object, the gravitational potential energy of a raised weight, and the elastic potential energy of a stretched spring. Electromagnetism reveals still other forms of energy. Among these are the energy transported by electromagnetic radiation. The electric field between the plates of a charged capacitor harbors energy as does a magnetic field. A charged battery stands ready to convert chemical energy into the energy of motion of electric charges.

Our notion of energy is unlikely to change very much from the elementary idea that energy measures the capacity to perform work. Our appreciation of energy grows as we find it thrust forward as a unifying concept in all areas of physics, indeed, throughout all science. Two primary attributes of energy quickly become apparent: Energy can be converted from one form to another, and energy is conserved. The first law of thermodynamics quantifies these attributes of energy. All forms of energy, mechanical, electrical, chemical, nuclear, and so on, are recognized together with the possibility of conversions from one type to another. In particular, the recognition of heat as a form of energy is necessary in order to preserve the conservation of energy. The most general statement of the first law of thermodynamics is this:

The total energy of the universe is constant.

Considerable analysis is needed to convert this sweeping statement into a working law of nature.

A mountain of *failures* stand as the experimental foundation of the first law. No one has succeeded in building a perpetual motion machine of the first kind—a device which produces more energy than it consumes. Many have tried, including crackpots and swindlers, who would hoax the gullible.

....

3.1.1 It is one thing to glibly pronounce that a perpetual motion machine will not work because it violates the first law of thermodynamics. It is quite another matter to understand precisely where a scheme goes wrong. Two proposed perpetual motion machines of the first kind are shown in Figure 3.1. Give a qualitative explanation of why they will not operate as suggested below. In other words, find the "fly in the ointment."

The machine in part (a) consists of a wheel whose axle is to act as the driveshaft of a vehicle. The spokes of the wheel are permanent magnets. The magnets are mounted so that their south poles all lie on the periphery of the wheel. Mounted above the center of the wheel is another permanent magnet. Symmetry considerations show that the net torque on the wheel is zero. To change this balance of torques, a magnetic shield is erected, so as to eliminate

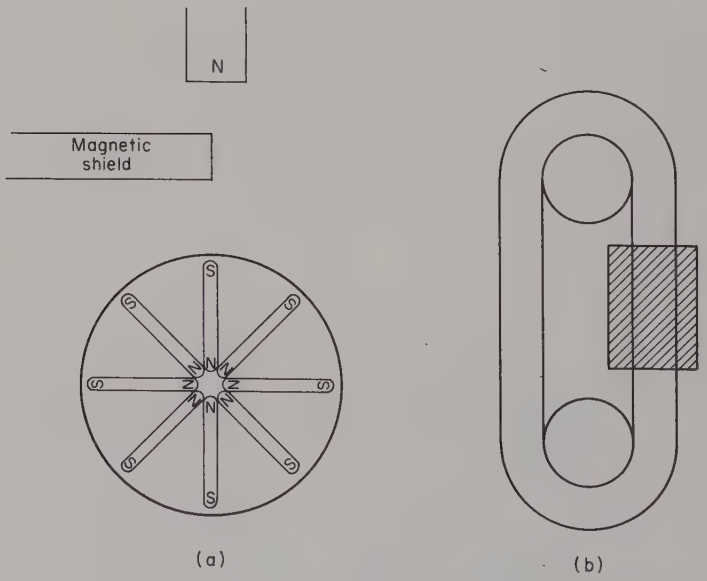


Figure 3.1. Two perpetual motion machines of the first kind: (a) magnetic wheel and (b) airtight tube.

the magnetic forces (and the torques) on the left side of the wheel. (A very crude magnetic shield might be simply a piece of soft iron. More elaborate magnetic shielding is used regularly, for example, to cancel the magnetic field of the earth in certain precision experiments.) With the shielding in position the net torque on the unshielded right side of the wheel should produce a counterclockwise rotation of the wheel.

The second device in part (b) consists of a flexible airtight tube designed to rotate around two pulley hubs. The axle of the lower hub serves as the driveshaft of a vehicle. Standing alone the gravitational force on the left side of the tube balances the force on the right, that is, there is no net torque and no rotation. By surrounding the right side of the tube with a "silo" filled with water a buoyant force is brought into play, that is, the portion of the tube inside the silo displaces water and is buoyed up by a force equal to the weight of the water displaced. With the buoyant force acting upward, the right side of the tube moves upward. Because of its continuous loop construction, a fixed volume of tubing is always submerged and the hubs rotate continuously, delivering power to the driveshaft.

•••••• ••••••

Example 3.1 The Sun, Energy Conversion, Nuclear Fusion, and the Bomb

The Sun serves as the source of energy by which life is maintained. The relentless stream of sunlight drives the winds, rains, and ocean currents and supplies energy for the complex processes of photosynthesis. The energy carried by sunlight is absorbed by green plants and used to convert carbon dioxide and water into oxygen and carbohydrates. Some of the absorbed energy sustains the plant, the remainder is converted into chemical energy—the energy associated with interatomic bonds. Green plants supply both the food we eat and the oxygen we breathe. The oxygen aids the release of chemical energy during metabolic processes.

Strangely enough the photosynthetic and metabolic processes which spend the energy delivered by sunlight are not understood as completely as the processes which release energy deep within the Sun. The source of stellar energy is nuclear *fusion*. At ordinary temperatures two hydrogen atoms can unite to form a hydrogen molecule. In the process they release energy to the surroundings. A similar sort of union, not of atoms but of atomic nuclei, releases energy in stellar interiors. The temperature in the core of the Sun is near 10 million kelvins. The primary nuclear fuel in the core is hydrogen, but it is *ionized* hydrogen. The high temperature there acts to strip away the atomic electron of the hydrogen atom. In most instances the hydrogen nucleus is a proton. Just as the union of two hydrogen atoms releases energy, the fusion of two protons is energetically favorable. Nuclear fusion results in the release of energy. The scale of the energy developed in fusion reactions is quite

substantial, about a million times that for the reaction between two hydrogen atoms.

According to Einstein's theory of relativity, mass is just another aspect of energy. Mass and energy are equivalent. The relation between mass m and energy E is given by that famous equation

$$E = mc^2$$

where c is the speed of light. In a reaction which releases energy ΔE, the masses of the reactants and products differ by Δm, where

$$\Delta E = \Delta m\, c^2$$

The form of the released energy is not determined by this equation, only its total amount. In fusion reactions, the released energy appears as kinetic energy of the products. Much of this in turn is converted into electromagnetic energy—sunlight—which then escapes from the solar surface.

The reactions between nuclei in stellar interiors are referred to as *thermonuclear* reactions. High temperature and density are very necessary conditions for achieving fusion reactions. In order for two protons to fuse they must approach one another very closely. Only when they approach within distances on the order of 10^{-13} cm can the attractive nuclear force effectively act to promote fusion. The attractive nuclear force is opposed by the electrical force of repulsion between the positively charged protons. In order to overcome this electrical repulsion the two protons must have a high relative kinetic energy as they approach. It is the high temperature which provides the kinetic energy, that is, the higher the temperature the faster the protons move. High density aids by making collisions between protons more frequent.

To date technological efforts to produce a controlled thermonuclear reaction have been unsuccessful. The high temperatures involved require that the reactants be confined by a magnetic field which keeps them away from the solid walls of the reactor vessel. Unfortunately, the magnetically confined configurations tend to be unstable. They quickly become turbulent and disperse. To date the only successful man-made thermonuclear reactions have been uncontrolled—they are better known as *hydrogen bombs*.

The explosion of TNT converts chemical energy into kinetic energy, heat, and light (radiant energy). A hydrogen bomb converts mass into other forms of energy, such as the kinetic energy of the reaction products and radiant energy. The hydrogen bomb is such an awesome energy converter that the energy release is reckoned in the equivalent number of millions of tons of TNT. The energy released by the explosion of one million tons of TNT is 4×10^{22} ergs. A unit of energy, the megaton, is defined as

$$1 \text{ megaton} \equiv 4 \times 10^{22} \text{ ergs}$$

The weight of a hydrogen bomb which releases 1 megaton of energy is presently a "secret."

One of the most valuable techniques a scientist can develop is the art of making order-of-magnitude estimates. These aid in deciding the feasibility of an idea. We illustrate the technique by making an order of magnitude estimate of the size of a hydrogen bomb.

The customary unit for expressing energies in nuclear physics is the million electron volt (MeV)

$$1 \text{ MeV} = 10^6 \text{ eV} = 1.6 \times 10^{-6} \text{ erg}$$

The mass of a proton is 1.67×10^{-24} gm. The mass energy of a proton $m_p c^2$ is 1.5×10^{-3} erg, or slightly less than 1000 MeV. The energy released by the fusion of two protons is approximately 1.4 MeV. In round numbers then nuclear fusion releases about 1/1000 of the mass energy of the reacting particles. It follows that the mass energy of a 1-megaton hydrogen bomb must be at least 1000 megaton, or 4×10^{25} ergs. Setting

$$mc^2 \simeq 4 \times 10^{25} \text{ ergs}$$

gives, for an estimate of the mass of the bomb,

$$m \simeq 4 \times 10^4 \text{ gm}$$

Scaling up, we can say that the mass of a 25-megaton hydrogen bomb is at least one million gm. A mass of 10^6 gm weighs very nearly 1 ton. We conclude that 25-megaton hydrogen bombs cannot be carried about in briefcases or vest pockets.

A hydrogen bomb is an uncontrolled thermonuclear fusion reaction. For over 25 years scientists have been striving to achieve a controlled fusion reaction. A fusion power reactor would have three distinct advantages over the now operative fission reactors. First, the fusion reaction leaves no radioactive ashes. In contrast, the by-products of fission reactions remain radioactive for many years and must be buried deep in the earth. Second, a good share of the energy released in the fusion reaction is in the form of kinetic energy of charged particles. By judiciously guiding these charges in a magnetic field it would be possible to induce currents in an external circuit. The fusion reactor would thereby permit the direct production of electrical power. In contrast, a fission reactor liberates energetic neutrons and leads to a less direct and less efficient production of electric power. The neutron kinetic energy is converted to heat, which is used to produce steam, which drives a turbine, which powers an electrical generator. The third point on which fusion reactors promise to outrank fission reactors concerns the nuclear fuel. The fissile isotopes of uranium and plutonium are scarce. Even with sophisticated "breeder" reactors which generate additional fissile material the projected

energy demands of the world will bankrupt our fission fuels in a few centuries. The basic fusion reactor fuel is deuterium.† Our supply of deuterium is virtually unlimited. We literally have oceans of it! About one out of every 6000 water molecules contains a deuterium atom, that is, it is DHO rather than H_2O. The recovery of deuterium from sea water is a relatively simple operation. Thus a supply of fusion fuel will be available as long as the oceans survive. ✧

....

3.1.2 The oxidation of 1 mole of sucrose ($C_{12}H_{22}O_{11}$) liberates 5.67×10^{10} ergs, in the form of heat.

(a) What is the decrease in mass of the system?

(b) What is the fractional mass to energy conversion ($\Delta m/m$ = mass converted/initial mass)?

3.1.3 In the year 1908, in central Siberia, a gigantic fireball crashed to Earth. The fireball was visible for many hundreds of kilometers. Strangely, no meteoric fragments were found. Within recent years it has been speculated that the fireball was composed of antimatter; a fragment of some antiplanet or anticomet, perhaps. A meeting of matter and antimatter is like a "kiss-of-death." They annihilate—their mass energy (mc^2) is converted into other forms of energy. For example, an antielectron (positron) interacts with an electron. The two particles disappear, annihilate, and in their place is an equivalent amount of electromagnetic energy in the form of γ rays. Antiprotons annihilate with protons, and so on. Most of the mass energy is quickly converted into heat and radiant energy.

In order to reach the ground, a chunk of antimatter would have to tunnel through our atmosphere. It would annihilate matter in its path and generate a monumental shock wave. (The Siberia shock wave made two transits around the earth.)

(a) Assume that a piece of antimatter travels vertically downward. Estimate the minimum mass which it must have in order to reach the surface of the earth.

(b) (Optional and a bit tricky) Treat the fireball as a sphere of uniform density. As it descends, annihilations reduce its radius. Show that the minimum

---

† There are three isotopes of hydrogen, differing in their nuclear composition. The nucleus of ordinary hydrogen consists of a single proton. Deuterium and tritium are the heavy isotopes. The nucleus of a deuterium atom consists of one proton and one neutron. The deuterium nucleus is called a *deuteron*. The heaviest isotope of hydrogen is tritium. The nucleus of a tritium atom is called a *triton* and contains one proton and two neutrons. The nuclear physicist is likely to designate these nuclei as p, d, and t or by the symbols ${}_1^1H$, ${}_1^2H$, and ${}_1^3H$. The atoms are often designated as H, D, and T.

initial radius which the fireball must have in order to reach the surface of the earth is

$$R = \frac{P}{4\rho_f g}$$

where P is the atmospheric pressure at the surface of the earth, g is the acceleration of gravity, and ρ_f is the mass density of the fireball.

(c) Compute the mass of the fireball of part (b), assuming it is composed of anti-iron ($\rho_f = 7.9$ gm/cm^3). Remembering that an equal mass of the atmosphere annihilates, compute the energy release, in megatons, resulting from the descent of such a fireball.

....

3.2 CALORIMETRIC DEFINITION OF HEAT

The concept of heat as a measurable quantity evolved during the 17th and 18th centuries. It was not until the middle of the 19th century that heat was recognized as a form of energy. The earliest quantitative theory viewed heat as a fluid, called caloric. This fluid was invisible and weightless. If the temperature of a body increased, it was because caloric had been absorbed. Cooling was accompanied by the loss of caloric. In order to explain certain calorimetry experiments it was necessary to postulate that caloric was conserved, that is, it could neither be created nor destroyed. The caloric theory was immensely successful. It could explain a great many thermal phenomena, such as the expansion of heated objects. The caloric theory was eventually discarded because it could not offer satisfactory explanations of certain phenomena. For instance, if two ice cubes are rubbed together in vacuum, they gradually melt. The caloric theory cannot adequately explain this phenomenon.

The language of thermodynamics contains many reminders of the caloric theory. Among these is the word "calorie," a unit of heat. The calorimetric definition of the unit of heat called the *calorie* is this:

> One calorie (1 cal) is the amount of heat required to raise the temperature of one gram of water from 14.5 °C to 15.5 °C.

A kilocalorie (1 kcal) is 1000 calories. In dietetics, the kilocalorie is generally referred to as the Calorie

$$1 \text{ kcal} \equiv 1 \text{ Cal} = 1000 \text{ cal}$$

Although heat was not originally interpreted as a form of energy, it was recognized as an extensive quantity—1 cal raises the temperature of 1 gm by 1 °C, but 5 cal are required to raise the temperature of 5 gm by 1 °C. For a

given mass of material it is also found that the amount of heat absorbed is directly proportional to the temperature increase. This direct proportionality holds true only for small changes of temperature and only so long as the heated substance does not undergo a change of phase, such as melting.

Let Q stand for the heat absorbed or rejected by a body of mass m and let ΔT denote its change in temperature. The observation that Q is proportional to $m\Delta T$ lets us insert a proportionality factor C and obtain the equation

$$Q = Cm\Delta T \tag{3.1}$$

The factor C is called the *specific heat capacity* and is a property of the physical and chemical composition of the material.† Equation (3.1) stands as a calorimetric definition of heat; but what is heat?! Equation (3.1) cannot answer. It tells us how to measure heat, not what heat measures.

....

3.2.1 If we consider 1 gm of a substance ($m = 1$ gm) the calorimetric definition of heat, Eq. (3.1), is

$$Q = C\Delta T$$

(a) Present an interpretation which justifies the designation of heat capacity for C.

(b) Present a second interpretation which reveals C as a measure of the thermal inertia of the material, its ability to resist a change of temperature. (Recall from mechanics that the mechanical inertia of a body, as measured by its mass, measures the object's resistance to acceleration.) You may wish to answer parts (a) and (b) by comparing two substances with different values of C.

3.2.2 One British thermal unit (Btu) is defined as the amount of heat required to raise the temperature of 1 lb of water by 1 °F. Show that 1 Btu equals approximately 252 cal.

3.2.3 Radio astronomers have provided us with a detailed picture of our own galaxy, the Milky Way, as well as glimpses of other distant galaxies. It has been estimated that the total energy received to date by all radio telescopes is less than 1 erg. To appreciate how minute an amount of energy this is, determine

(a) the temperature increase of 1 gm of water when it receives 1 erg of heat energy;

---

† Where no confusion can result we often shorten "specific heat capacity" to "specific heat."

(b) the height to which you could raise an ant having a mass of 40 mg by expending work of 1 erg.

....

At ordinary temperatures, the specific heats of most solids and liquids vary only slightly with pressure and volume. For gases it is much more important to distinguish the conditions under which heat is absorbed or rejected. For example, the specific heat of a gas maintained at a constant pressure is greater than the specific heat measured when the gas is held at a constant volume. It was this curious fact concerning the specific heats of gases which led Robert Mayer to first hypothesize that heat was a form of energy and that it could be converted into other forms of energy. For this singular stroke of genius, Mayer shares (with James Prescott Joule) the credit for formulating the first law of thermodynamics. As a prelude to stating the first law in equation form the important concept of work is reviewed in the next section. We conclude this section with a brief survey of calorimetry, the science of heat measurement.

The basic question in calorimetry is this: If two objects, initially at different temperatures, are placed in thermal contact and permitted to reach equilibrium, what is their final temperature? As a specific example, consider a situation in which a 500-gm block of aluminum at a temperature of 400 K is placed in a tank containing 1000 gm of water at a temperature of 300 K. The specific heat capacities of aluminum and water are 0.215 cal/gm-°C and 1.000 cal/gm-°C, respectively. Let T_f denote the final temperature. From (3.1) we find the heat rejected by the aluminum to be

$$Q_A = m_A C_A \Delta T_A = 107.5(T_f - 400\,\mathrm{K}) \quad \text{cal}$$

The heat absorbed by the water is

$$Q_W = m_W C_W \Delta T_W = 1000(T_f - 300\,\mathrm{K}) \quad \text{cal}$$

The magnitudes of Q_A and Q_W are equal, that is, the heat rejected by the aluminum is absorbed by the water. However, Q_A is negative (heat rejected) and Q_W is positive (heat absorbed). We can therefore write $Q_A + Q_W = 0$ and solve for T_f. The result is $T_f \simeq 310\,\mathrm{K}$.

The method is readily generalized. Thus, if we have N substances, with masses $m_1, m_2, \ldots, m_N$, initially at temperatures $T_1, T_2, \ldots, T_N$, we can write

$$\sum_{j=1}^{N} Q_j = 0, \qquad Q_j = m_j C_j (T_f - T_j)$$

thereby obtaining

$$T_f = \frac{\sum_{j=1}^{N} m_j C_j T_j}{\sum_{j=1}^{N} m_j C_j}$$

Problem 3.2.4 indicates how the method of mixtures can be used to measure heat capacities.

....

3.2.4 The method of mixtures is a basic calorimetric technique for determining specific heat capacities. It relies on the ideas presented above. However, the unknown quantity is one of the specific heat capacities. The final temperature is an experimental datum.

(a) Show that the specific heat capacity of the component labeled 1 is given by

$$C_1 = \frac{\sum_{j=2}^{N} m_j C_j (T_f - T_j)}{m_1 (T_1 - T_f)}$$

(b) Place 600 gm of lead (Pb) shot in an aluminum calorimeter, which holds 100 gm of water. The water and calorimeter are initially at a temperature of 80 °C. The initial temperature of the lead is 25 °C. If the final temperature of the mixture is 72 °C, determine the specific heat of lead. The mass of the calorimeter is 20 gm and its specific heat capacity is 0.215 cal/gm-°C.

3.2.5 The air temperature above coastal areas is profoundly influenced by the large specific heat of water (1 cal/gm-°C). One reason is that the heat released when 1 cm^3 of water cools by 1 °C will heat an enormously greater volume of air by this same amount. Estimate the volume of air which can be heated by 1 °C by the 1 cal released by the cooling of 1 cm^3 of water by 1 °C. The specific heat of air is approximately $\frac{1}{4}$ cal/gm-°C. The density of air under standard conditions is 0.00129 gm/cm^3.

....

3.3 WORK IN THERMODYNAMICS

The concepts of work and energy are related closely. It will be recalled from mechanics that a force **F**, acting through a displacement **ds**, performs an amount of work dW given by the scalar product of **F** and **ds**,

$$dW = \mathbf{F} \cdot \mathbf{ds} \tag{3.2}$$

To determine the total work performed over a finite displacement, the differential elements of work must be summed, that is, integrated. We can write, for the total work,

$$W = \int_A^B dW = \int_A^B \mathbf{F} \cdot \mathbf{ds} \tag{3.2a}$$

The symbolism $\int_A^B dW$ means "add up all elements of work performed between

A and B." In order to evaluate the integral it is generally necessary to specify the path from A to B. We say that W is "path-dependent."

•••••• ••••••

3.3.1 An object moves in the x–y plane under the influence of a force which has only an x component

$$F_x = xy$$

The work done by this force acting along a path leading from the origin to the point $x = 1$, $y = 1$ depends on the nature of the path. Demonstrate this fact by showing that the work done along the straight-line path $y = x$ is not the same as the work done along the parabolic path $y = x^2$.

•••••• ••••••

The mechanical work of a force is the most familiar form of work. The work performed on a thermodynamic system may take other forms. In thermodynamics, *work* is broadened to include *all forms of energy transfer excepting heat.*† Suppose, for example, that the system is an electric cell—a battery. While the cell is being charged, electrical work is being done by an external electromotive force (emf). This work is stored within the cell as internal energy. In our study of thermodynamics we shall be particularly concerned with two forms of work. One is the work performed by a fluid as it expands or contracts. The other is the work done in altering the magnetic structure of certain types of solids.

Work of expansion and magnetic work are studied in depth in the next chapter. For now it suffices to recognize that thermodynamic systems can perform work of various forms. There are situations where it is more natural to envision the surroundings as performing work on the system. This can be described by saying that the system performs a "negative amount of work." We shall use W (or dW) to designate work performed *by* the system against its surroundings. This amounts to nothing more than choosing a name for the work done by the system and W is the name! In some instances W may be negative. For example, an expanding gas performs a positive amount of work against its surroundings. If a gas is compressed, it does negative work, that is, the surroundings do a positive amount of work. It should therefore come as no surprise when a calculation produces a negative value of W.

•••••• ••••••

3.3.2 If W is the name for "work done by the system against its surroundings," what is $-W$ the name for?

•••••• ••••••

† This statement refers to *closed* systems. For open systems the transfer of matter also results in the transfer of energy.

3.4 INTERNAL ENERGY

Consider a system which is thermally isolated from its surroundings. The jargon of thermodynamics describes this situation by saying that the system is "surrounded by adiabatic walls." If the system is insulated, it can perform work only at the expense of its energy, but what, precisely, is meant by "the energy of the system"? If the surroundings perform a positive amount of work on the system, we instinctively feel that the energy of the system has been increased. This intuitive feeling can be parlayed into an operational definition of internal energy. Let W denote the work done by the system during a process which carries it from some initial equilibrium state to some final equilibrium state. Let U designate the internal energy of the system—the quantity we wish to define operationally. An appeal to the conservation of energy lets us write

Internal energy of the system in its initial state	=	Internal energy of the system in its final state	+	Work done by the system on its surroundings

In symbolic form this becomes

$$U_i = U_f + W$$

where the subscripts refer to the initial and final states. We already have an operational definition of work. This relation therefore serves to operationally define the *change* in the internal energy of the system. By rearranging terms it may be written

$$U_f - U_i = -W \tag{3.3}$$

where $U_f - U_i$ is the change in the internal energy of the system and $-W$ is the work done *on* the system by its surroundings. Equation (3.3) is often referred to as the adiabatic form of the first law of thermodynamics.

Two remarks on (3.3) are worthy of mention. First, only the difference $U_f - U_i$ is operationally defined. It is permissible to select some particular state and arbitrarily assign it the value of zero internal energy. The *standard state* of an element or compound is defined as its most stable physical form at a pressure of one atmosphere and a temperature of 25 °C ($\simeq$298 K). For reasons to be given in Chapter 8 the value zero is assigned not to the internal energy of the standard state, but rather to the quantity (P is pressure, V is volume)

$$H \equiv U + PV$$

known as the *enthalpy*. By definition then, the enthalpies of 1 liter of gaseous helium, 1 kg of liquid mercury, and 1 pinch of table salt are all zero at a pressure of 1 atm and a temperature of 25 °C.

The second comment bears on the nature of internal energy. A definite value of U is associated with each equilibrium state of a system. For this

reason U is referred to as a *state function*. The value of U for a given state is independent of the previous history of the system. Stated otherwise, the thermodynamic "path" leading to a given state does not affect the value of U for that state. In this respect, internal energy is analogous to wealth. A man who is a millionaire today may have been born in poverty or with a silver spoon in his mouth. He may have amassed and lost several fortunes. He may have inherited the money only the day before. Which, if any, of these financial paths he followed is irrelevant in so far as his present wealth is concerned.

Our definition of internal energy, though operational, may not leave us with much of a feel for the quantity. Our previous experience suggests that U be regarded as the sum of the kinetic and potential energies of the parts—the atoms and molecules which constitute the microscopic structure of the system. While this is a healthy viewpoint it is not thrust upon us by thermodynamics. Thermodynamics operates at the macroscopic level. If we want to see inside a system and understand what makes it tick, we must erect a microscopic model and a theory to go with it. We must believe in atoms and puzzle over the forces which bind them together to form molecules and solid objects. The disappointment of learning that thermodynamics cannot answer certain questions inspired men like Rudolph Clausius, James Clerk Maxwell, Ludwig Boltzmann, and Josiah Willard Gibbs to develop the kinetic theory of gases and statistical mechanics. These two disciplines, which permit events to be followed at the microscopic level in a statistical or average sort of fashion, are developed in the second half of this text.

....

3.4.1 Gases are notoriously poor heat conductors. For this reason the rapid expansion or compression of a gas may be regarded as an adiabatic process, that is, it is over before the gas has a chance to absorb or reject a significant amount of heat.

A man exerts a constant force of 1000 N in compressing the gas in a pump cylinder. The force acts over a distance of 50 cm.

(a) How much work is done on the gas?

(b) What is the change in the internal energy of the gas? What assumptions have you made in arriving at your answer?

....

3.5 THE FIRST LAW OF THERMODYNAMICS

Credit for formulating the first law of thermodynamics is shared by Robert Mayer and James Prescott Joule. Working independently and in response to quite different motivations, both arrived at the same conclusions, to wit:

(a) Heat is one of many different forms of energy.
(b) Energy is conserved—it may be transformed, but not created or destroyed.

In 1842 Mayer presented the first determination of the mechanical equivalent of heat, the amount of mechanical energy equivalent to 1 cal. Mayer's value was within 20% of the presently accepted value of

$$1 \text{ cal} = 4.186 \text{ joule} = 4.186 \times 10^7 \text{ erg} = 0.0040 \text{ Btu}$$

James Prescott Joule, an English brewer, based his belief in the conservation of energy on religious grounds,† "... the grand agents of nature are, by the Creator's fiat, indestructible; ... (1843, p. 435)." In a series of experiments spanning 40 years Joule measured the mechanical equivalent of heat. He demonstrated that the dissipation of a given amount of mechanical or electrical energy always resulted in the evolution of a proportional amount of heat.

The general form of the first law of thermodynamics recognizes that a change in the internal energy of a system can be accomplished by the performance of work, by the exchange of heat, or by a combination of both. Let Q denote the quantity of heat added to a system during a process which carries it from some initial equilibrium state to some final equilibrium state. In general such a process changes the internal energy of the system and results in the performance of work. The adiabatic form of the first law, Equation (3.3), may be generalized to take account of heat exchange by stating that

$$\begin{Bmatrix}\text{heat added}\\ \text{to a system}\end{Bmatrix} \text{ appears as } \begin{Bmatrix}\text{a change in the}\\ \text{internal energy}\end{Bmatrix} \text{ and/or } \begin{Bmatrix}\text{is used to}\\ \text{do work}\end{Bmatrix}.$$

The equation expressing this statement, the general form of the first law of thermodynamics, is

$$Q = U_f - U_i + W \tag{3.4}$$

We remark that Q is the *name* for the net amount of heat added to the system. If Q should turn out to be negative, it simply means that the system rejected a net amount of heat. In an adiabatic process $Q = 0$ and (3.4) reverts to the adiabatic form (3.3).

Example 3.2

A system absorbs 1.2×10^6 cal of heat and performs 1.7×10^{14} ergs of work against its surroundings. Determine the change in the internal energy of the system.

---

† Joule's accounts of his first determination of the mechanical equivalent of heat were published in the *Philosophical Magazine, Series 3*, **23**, 263, 347, 435 (1843).

We rewrite (3.4) as

$$U_f - U_i = Q - W$$

Since the heat is absorbed by the system,

$$Q = +1.2 \times 10^6 \text{ cal}$$

Positive work is done by the system against the surroundings

$$W = +1.7 \times 10^{14} \text{ ergs}$$

If we elect to express $U_f - U_i$ in ergs, Q must be converted to ergs. The student should verify that this conversion gives

$$Q = 0.50 \times 10^{14} \text{ ergs}$$

Inserting these values of Q and W into the first law gives

$$U_f - U_i = -1.2 \times 10^{14} \text{ ergs}$$

The minus sign means that the final internal energy is less than the initial by 1.2×10^{14} ergs. Internal energy is sacrificed because the work performed exceeded the heat absorbed. ✧

....

3.5.1 The surroundings perform 1.7×10^{14} ergs of work on the system of Example 3.2. If the system absorbs 1.2×10^6 cal of heat, compute the change in its internal energy.

3.5.2 The Joule constant, $J = 4.186$ joule/cal, is often referred to as the "mechanical equivalent of heat." At the end of his 1842 essay Mayer expresses his results for the mechanical equivalent of heat by stating that the warming of a given weight of water from 0 °C to 1 °C is equivalent to the mechanical energy acquired by the same weight of water in falling from a height of 365 m. Show that the value of J for these data is approximately

$$J = 3.58 \times 10^7 \text{ ergs/cal}$$

Assume that the specific heat of water is 1 cal/gm-°C and take the acceleration of gravity to be 981 cm/sec^2.

....

In Section 3.3 it was observed that work is generally a path-dependent quantity. In Section 3.4 it was emphasized that internal energy is a property of the system. The value of U for a particular thermodynamic state is independent of the prior history of the system. Therefore, the change $U_f - U_i$ depends solely on the initial and final states. It is independent of the thermodynamic path joining the two states. It now follows from the first law that Q is also a path-dependent quantity, that is, Q is the sum of a path-independent quantity $U_f - U_i$ and a path-dependent quantity W. The path dependence of

Q follows still more directly from the first law if we consider a *cyclic process*. In a cyclic process the initial and final states are the same so that $U_f = U_i$, and the first law reduces to

$$Q = W \tag{3.4a}$$

that is, it states that the total work done by the system over the cycle equals the net heat absorbed by the system. Since W depends on the nature of the cyclic path the heat Q is likewise path-dependent.

•••••• ••••••

3.5.3 For many years the Patent Office was deluged with applications for patents on various types of perpetual motion machines. This flood subsided abruptly and completely when the rules were amended to require that a working model accompany each application. Suppose that you were employed in the Patent Office prior to this ruling and that an application for a patent on a heat engine was presented to you, the performance figures claimed for the engine being as follows:

Net heat consumption per cycle: 10 kcal
Mechanical work output per cycle: 10^{12} ergs
Electrical work output per cycle: 10^5 watt-sec

Would it be necessary to give detailed consideration to the application? Explain.

•••••• ••••••

In situations where a thermodynamic process proceeds smoothly we can safely envision it as a continuous sequence of small changes. Mathematically we can represent the quantities appearing in the first law by infinitesimals. The miniaturized form of the first law is

$$dQ = dU + dW \tag{3.5}$$

Of course, the physical content of the law is unaltered by rewriting it in this form. Furthermore, this form presents a pitfall of which one must be aware. This concerns the path dependence of heat and work. The fact that we represent heat and work as dQ and dW does not imply the existence of state functions, Q and W, which measure the heat and work content of a system. Thus dQ and dW are mathematically infinitesimal, but they are not true differentials†; they denote small amounts of heat and work.

In order to fully exploit the rich content of the first law we must determine under what conditions it is permissible to express dQ, dU, and dW in terms of the thermodynamic coordinates (P, V, T, and so on), their changes (dP, dV,

† Some authors use a special notation such as $đW$ or δQ as a reminder of this pitfall.

dT, and so on), and other measurable properties, such as the specific heat capacities. Such a description is possible for processes which qualify as reversible. The concept of reversibility is discussed in the next chapter. Also in the next chapter, we develop expressions for dU and dW in terms of the thermodynamic coordinates and their changes. The final step in the program, expressing dQ in a similar fashion, carries us to the second law of thermodynamics.

REFERENCES

A beneficial use of hydrogen bombs, for the propulsion of interstellar vehicles, is discussed in:

F. J. Dyson, Interstellar Transport, *Phys. Today* **21**, 41 (October 1968).

The arguments against building such a spacecraft are related in:

F. J. Dyson, Death of a Project, *Sci.* **149**, 141 (1965).

The multifaceted nature of energy is surveyed in spectacular fashion in:

M. Wilson, "Energy." Time Inc., New York (1963).

The importance of energy to all technologies is brought home in:

F. J. Dyson, The Search for Extraterrestrial Technology, *in* R. Marshak (ed.), "Perspectives in Modern Physics." Wiley (Interscience), New York (1966).

An entertaining account of attempts to violate the first and second laws of thermodynamics is given in:

S. W. Angrist, Perpetual Motion Machines, *Sci. Am.* **218**, 115 (Jan. 1968).

See also the "dynamaforce generator" described by:

M. Gardner, Mathematical Games. *Sci. Am.* **226**, 100 (Feb. 1972).

A convenient source of biographical data on Joule, Mayer, and other pioneers is:

W. F. Magie, "A Source Book in Physics." McGraw-Hill, New York (1935).

The caloric theory of heat died a lingering death. See:

S. C. Brown, The Caloric Theory of Heat, *Am. J. Phys.* **18**, 367 (1950).

Professor Brown is the foremost authority on the life and works of Count Rumford, a remarkable individual who sparked the rise of the kinetic view of heat. See:

S. J. Brown, "Count Rumford, Physicist Extraordinary." Doubleday, Garden City, New York (1962).

Discussions of vector algebra and path integrals are given by:

M. H. Hull, Jr., "The Calculus of Physics," Chapters 3 and 4. Benjamin, New York (1969).

Chapter 4

Thermodynamic Processes and Thermodynamic Work

In this chapter the important concepts of quasi-static and reversible processes are developed. The notion of thermodynamic work is introduced and illustrated in terms of the work performed by an expanding fluid and the work performed when a material is subjected to a magnetic field. The chapter concludes with an exploration of the relationship between specific heat capacities and the internal energy.

4.1 THERMODYNAMIC PROCESSES

A thermodynamic process results in changes in the thermodynamic coordinates and/or changes in the surroundings.† As examples of thermodynamic processes we may cite:

(a) a chemical reaction in which a precipitate forms;
(b) the explosive combination of hydrogen and oxygen to form water;
(c) the diffusion of perfume vapor—an effect observed quite often by girl watchers;

† A cyclic process leaves the state of the system unchanged; all thermodynamic coordinates return to their original values.

(d) the compression of a fluid by a piston;
(e) the stretching of a piece of wire;
(f) the magnetization of a chromium potassium alum pill.

The first three processes are irreversible. Irreversible processes occur spontaneously within a system which is not in thermodynamic equilibrium. Heat flow is one type of irreversible process. Heat flows spontaneously, that is, without any outside help from regions of higher temperature to regions of lower temperature. Experience reveals that such processes tend to produce a state of mutual equilibrium between the system and its surroundings. Once attained, an equilibrium state persists as long as the surroundings remain unchanged. The designation irreversible is most apt, for such processes act in an undirectional fashion—to drive a system toward an equilibrium state. The study of irreversible processes properly belongs to the field of non-equilibrium thermodynamics. Equilibrium thermodynamics is concerned with quasi-static reversible processes.

4.2 QUASI-STATIC PROCESSES

A quasi-static process is one in which the surroundings and thermodynamic coordinates of the system change slowly—so slowly, that the process may be viewed as one in which the system passes through a succession of equilibrium states. Geometrically, a quasi-static process may be represented by a path which joins the initial and final equilibrium states by a succession of intermediate equilibrium states. If the equation of state can be represented by a real surface (for example, the P–V–T surface of an ideal gas), the path of the quasi-static process can actually be traced on this surface.

Among the many quasi-static processes which a thermodynamic system can undergo, members of the "iso" family are favorites, for example,

$$\text{Isothermal:} \quad T = \text{constant}$$

$$\text{Isobaric:} \quad P = \text{constant}$$

$$\text{Isovolumic:} \quad V = \text{constant}$$

Regardless of the nature of a quasi-static process it is important to distinguish between the equation which specifies the path of the process and the equation of state. Thus, for example, an ideal gas obeys the equation of state $PV = nRT$ at every stage of a quasi-static process whether the process is isobaric or isothermal, that is, regardless of the path. The equation of state defines the nature of the thermodynamic *surface*. The particular quasi-static process

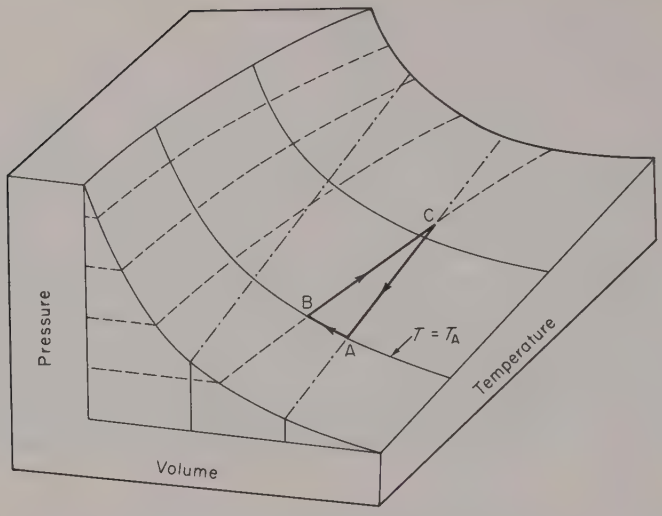


Figure 4.1. The *P–V–T* surface for an ideal gas. The path from A to B describes an isothermal (constant temperature) compression of the gas. The path from B to C describes an isobaric (constant pressure) expansion of the gas. The path from C to A describes an isovolumic (constant volume) decompression.

defines a *path on that surface*. Figure 4.1 shows the thermodynamic surface for the ideal gas. The three paths AB, BC, and CA describe isothermal, isobaric, and isovolumic processes, respectively.

It is frequently desirable to represent a *P–V–T* surface in two dimensions. This is usually done by projecting the *P–V–T* surface onto the *P–V* plane. Figure 4.2 shows the *P–V* diagram for the ideal gas surface of Figure 4.1.

....

4.2.1 Figure 4.2 shows the projection of the *P–V–T* surface of Figure 4.1. The projection of Figure 4.2 is onto the *P–V* plane. Make sketches of the projections onto the *P–T* and *V–T* planes. Include lines showing the path of the cyclic process $A \rightarrow B \rightarrow C \rightarrow A$.

....

The importance of quasi-static processes is understandable. We wish to describe processes in terms of the thermodynamic coordinates of the system, and these are well defined only for equilibrium states. A semiquantitative criterion for quasi-static behavior may be stated as follows: Let τ_p denote a time characteristic of the process; for example, the time required for the

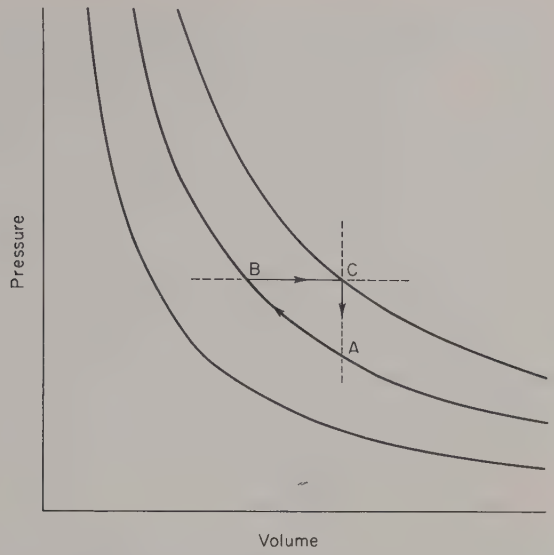


Figure 4.2. The *P–V* plane projection of the ideal gas surface. The cyclic process, A→B→C→A, is the same as that indicated in Figure 4.1.

thermodynamic coordinates to undergo a measurable change. The rate at which the system approaches equilibrium is also characterized by a time, which we denote by τ_E. For a process to qualify as quasi-static we demand

$$\tau_E \ll \tau_p \tag{4.1}$$

that is, the time needed to reach equilibrium must be small compared to the time required for the system to change its thermodynamic state. The condition $\tau_E \ll \tau_p$ means that the irreversible processes work fast, always keeping the system in near-equilibrium states. For example, consider a process in which a piece of metal is slowly heated. Heat conduction is the irreversible process which strives to achieve temperature equality throughout the metal. If the heat is added very slowly, by maintaining the heating element at a temperature just slightly above that of the metal, the temperature will rise uniformly throughout the block. In other words, thermometers positioned at different locations in the metal would give identical readings if heat conduction works fast in comparison to the rate of heating. For heat conduction, τ_E is the time for a temperature change to spread throughout the block. The time required to add enough heat to produce a detectable change in temperature is τ_p. Rapid heat conduction and the slow addition of heat cooperate to make $\tau_E \ll \tau_p$ and make the process quasi-static.

4.3 REVERSIBLE PROCESSES AND THERMODYNAMIC WORK

The example of a gas being compressed quasi-statically by a piston serves admirably to introduce the notions of a *reversible process* and *thermodynamic work*. Figure 4.3 suggests how a quasi-static compression might be accomplished. Consider first the behavior of the system when there is no friction between the piston and the cylinder walls. Adding or removing a grain of sand is to be regarded as producing an infinitesimal change in the surroundings. At any stage we can halt the compression and initiate expansion by removing a grain of sand, provided there is no friction. If the sand is removed grain by grain, so as to reverse the order in which it was added, the gas will retrace the thermodynamic path it followed during the compression. A significant point is that the reversal can be initiated by an infinitesimal change in the surroundings.

The quasi-static frictionless compression just described is an example of a *reversible process*. A general definition is this:

> A reversible process is a quasi-static process which can be reversed by an infinitesimal change in the surroundings.

Please note that a reversible process is not necessarily one which goes one way then back again. Reversible implies it *can* be reversed.

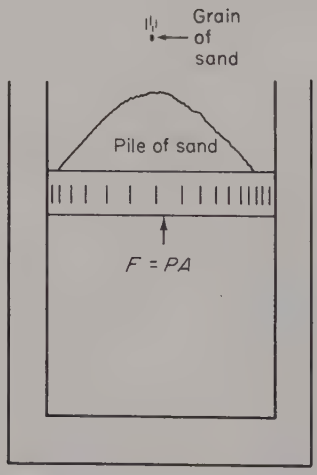


Figure 4.3. An example of the quasi-static compression of a gas. This can be achieved by slowly piling sand on the piston. If the piston moves without friction the removal of one grain of sand (an infinitesimal change) causes the piston to reverse its motion. If there is friction present, the removal of one grain need not produce a reversal of the piston.

The presence of friction is fatal to reversibility. If there is friction, we can still compress the gas quasi-statically by adding sand slowly but—and here is where the reversibility is lost—the compression cannot be reversed by an infinitesimal change in the surroundings. Removing one grain of sand—an infinitesimal change—will not reverse the process. It requires a finite change in the surroundings to revert from compression to expansion.

It is, of course, not possible to eliminate friction. Nevertheless, it often is possible to render friction negligible by comparison with the force exerted by the gas. Very frequently one deals with a solid or a liquid whose surface expands against the surrounding atmosphere. Under such circumstances the motion of the piston, that is, the surface, is as nearly frictionless as one can imagine.

An important consequence of reversibility is this: The work done by the system can be expressed in terms of its thermodynamic variables and their changes. We illustrate this point using the gas–piston system. If P is the gas pressure and A the cross-sectional area of the piston, the force exerted on the piston by the gas is

$$F = PA \tag{4.2}$$

Let the piston move outward (expansion) a distance ds. The work done by the gas is

$$dW = F\,ds = PA\,ds \tag{4.3}$$

However, $A\,ds = dV$ is the change in the volume of the gas. Thus, the work done by the gas during a quasi-static change of volume is

$$dW = P\,dV \tag{4.4}$$

The expression (4.4) is so simple that its most remarkable characteristic may escape our notice. We have succeeded in expressing the work solely in terms of the thermodynamic coordinates of the system. The nature of the external force and other characteristics of the surroundings do not appear in the expression for dW. This remarkable result is possible only to the extent that the compression is quasi-static and friction-free. Friction would introduce an additional term in (4.4) representing the work done by frictional forces. The work $dW = P\,dV$ is often referred to as *thermodynamic work*, as it is expressed entirely in terms of the thermodynamic coordinates of the system. By contrast, the work done by a frictional force would not be classified as thermodynamic work because such a force is not a thermodynamic coordinate of a system. As only thermodynamic work will be of much concern we generally do not retain the over worked adjective "thermodynamic" in referring to such work.

In Chapter 3, we called attention to the path dependence of work and heat. Having an explicit expression for the work of expansion

$$dW = P\,dV$$

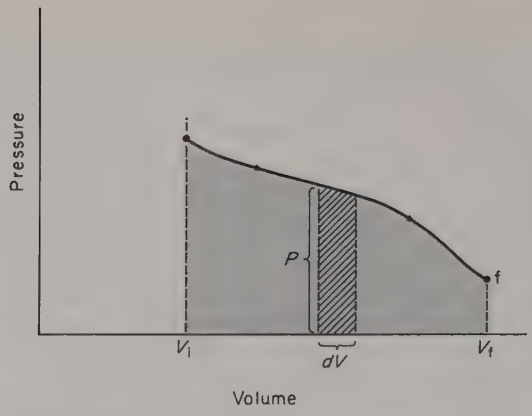


Figure 4.4. Geometrically, the work $dW = P\,dV$ corresponds to the hatched area. The total work done by the gas in the expansion from V_i to V_f corresponds to the total area beneath the P–V path.

affords us the opportunity to illustrate the path dependence of thermodynamic work. The work done by the system during a finite change of volume is found by integrating $dW = P\,dV$. We write

$$W = \int_i^f P\,dV \tag{4.5}$$

with i and f denoting the initial and final states, respectively. A geometric representation of the integral makes it clear that W depends on the thermodynamic path between states. Geometrically W corresponds to the shaded area of Figure 4.4. This area clearly depends on the nature of the thermodynamic path connecting the initial and final states.

Example 4.1 The Path Dependence of Work

To illustrate the path dependence of W we calculate the thermodynamic work done during the expansion of an ideal gas. Two types of expansion (that is, two different paths) are considered. Figure 4.5 shows the two paths. The first path consists of an isobaric expansion at a pressure P_1 followed by an isovolumic decompression at volume V_2. The other path considered is the isothermal expansion along the isotherm $T = T_1$. For the isobaric expansion

$$W_{\text{isob}} = \int_{V_1}^{V_2} P_1\,dV = P_1(V_2 - V_1) \tag{4.6}$$

where isob is isobaric. There is no work done during the isovolumic decompression, so that the total work for the path 1→A→2 is

$$W_{1A2} = P_1(V_2 - V_1)$$

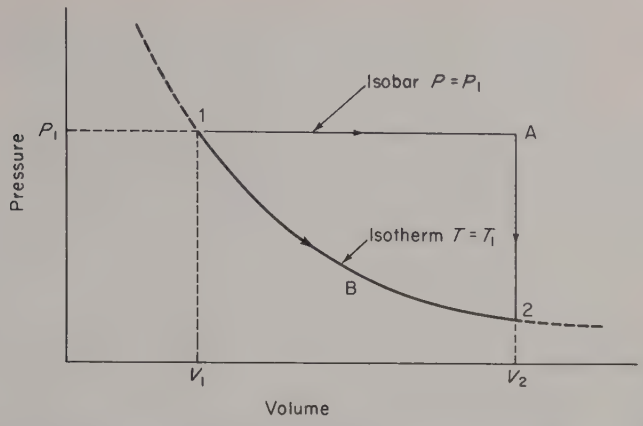


Figure 4.5. Two different paths connecting the initial and final states. The work for the path $1 \rightarrow A \rightarrow 2$ exceeds that for the isothermal path $1 \rightarrow B \rightarrow 2$.

A check of Figure 4.5 verifies that this work is represented geometrically by the area of the rectangle of height P_1, and length $V_2 - V_1$. Consider next the work performed by the gas during an isothermal expansion. With $T = T_1$ throughout the expansion the pressure is related to the volume by

$$P = \frac{RT_1}{V}$$

The isothermal work is†

$$W_{1B2} = RT_1 \int_{V_1}^{V_2} \frac{dV}{V} = RT_1 \ln\left(\frac{V_2}{V_1}\right) \tag{4.7}$$

Geometrically the work W_{1B2} corresponds to the area beneath the isothermal path $1 \rightarrow B \rightarrow 2$ in Figure 4.5.

It is obvious from the inspection of Figure 4.5 that $W_{1A2} > W_{1B2}$ even though both paths connect the same initial and same final states. The general conclusion to be drawn from this specific example is that the *thermodynamic work depends on the path of the process* connecting the initial and final states. ✧

One other aspect of the geometrical interpretation of thermodynamic work should be noted. A cyclic process may result in the performance of net thermodynamic work by the system. Consider, for example, a process which takes 1 mole of an ideal gas through the cycle (compare Figure 4.5) $1 \rightarrow A \rightarrow 2 \rightarrow B \rightarrow 1$.

---

† Recall that $\int_a^b dx/x = \ln(b/a)$, where $\ln x$ refers to the natural logarithmic base.

The total work done by the system over the cycle, abbreviated cyc, is

$$W_{\text{cyc}} = W_{1A2} + W_{2B1} = P_1(V_2 - V_1) - RT_1 \ln\left(\frac{V_2}{V_1}\right)$$

Geometrically this work is represented by the net area enclosed by the path of the cycle on the P–V diagram. This result is generally valid: the quasi-static work done by the pressure forces over a full cycle can be written

$$W = \oint P\,dV \tag{4.8}$$

The symbolism $\oint$ signifies that the integration is over a full cycle. Geometrically, $\oint P\,dV$ represents the area enclosed by the path describing the cycle on a P–V diagram. Figure 4.6a–d illustrates this point. During expansion the volume increases so that dV and thus $dW = P\,dV$ are positive. The total work

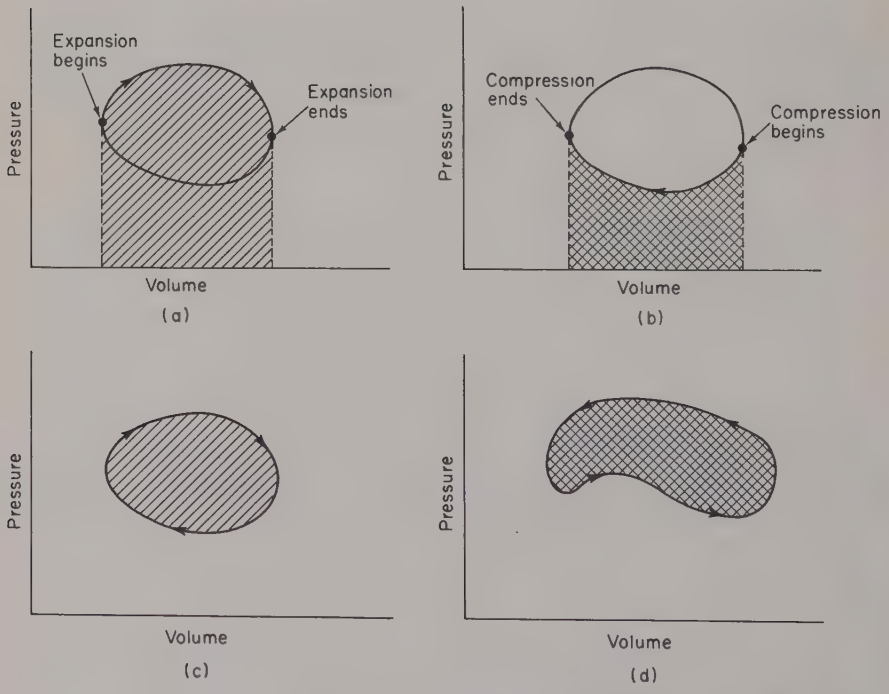


Figure 4.6. (a) The work of expansion is positive; it corresponds to the hatched area. (b) The work of compression is negative; its magnitude corresponds to the cross-hatched area. (c) The net work is positive. It corresponds to the area enclosed by the clockwise path. (d) A cycle for which the net work done is negative; its magnitude corresponds to the enclosed area. Note that the cycle generates a counterclockwise path.

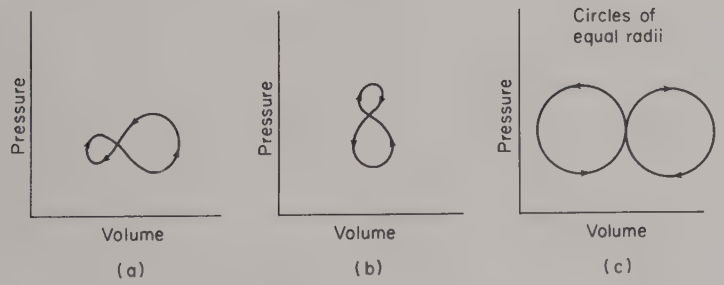
of expansion is thus positive and its magnitude corresponds to the hatched area of Figure 4.6a. The compression from V_2 to V_1 reduces the volume so that dV and dW are negative. The total work of compression is negative but its magnitude is given by the cross-hatched area of Figure 4.6b. The net work done by the pressure force over a full cycle is evidently the sum of the works of expansion and compression. Geometrically the net work corresponds to the difference between the corresponding areas. Figure 4.6c shows that this is just the area enclosed by the path of the cycle. In this case the net work done by the system is positive, the pressure during expansion being greater than that during compression. Note that the cyclic process generates a clockwise path on the P–V diagram. A clockwise path on a P–V diagram always corresponds to positive work.

Figure 4.6d shows a cyclic process for which the net work is negative. The work of compression is negative and its magnitude exceeds the positive work of expansion. Note that the cycle traces out a counterclockwise path, which on a P–V diagram always corresponds to negative work. The magnitude of the net work still corresponds to the enclosed area.

In the next section we study magnetic work and show how it may be expressed in a form analogous to (4.4).

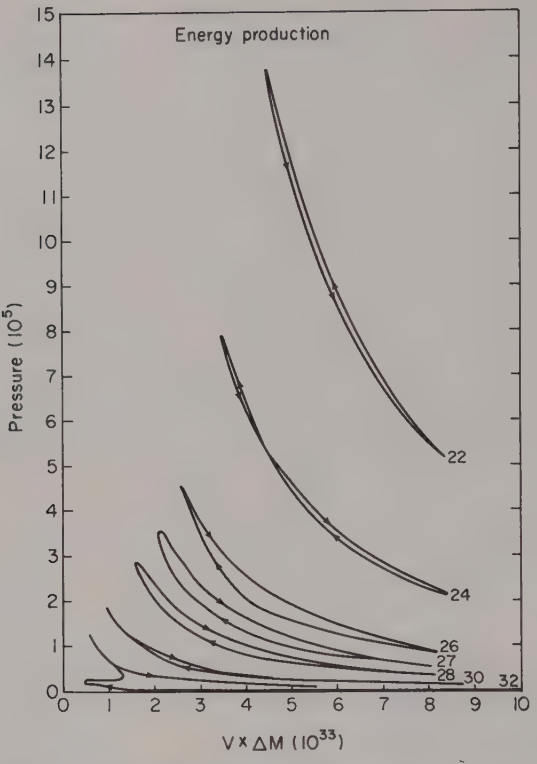
•••••• ••••••

4.3.1 The paths for two cyclic processes are indicated in the accompanying figure. Indicate for each part whether the net work done is positive, negative, or zero. Explain the basis for your choice.



4.3.2 Certain classes of stars show periodic variations of their luminosity. This behavior has been traced to periodic pulsations of the star. A detailed study of the pulsations of such stars has been made by Robert F. Christy.† In order to follow the temporal evolution of the pulsations, Christy treats the star as a set of concentric shells of mass. He solves a set of nonlinear equations

† R. F. Christy, Lectures on Variable Stars, delivered at the *3rd Summer Inst. Astronomy and Astrophysics, Stony Brook, New York (1969)*. See also *Rev. Mod. Phys.* **36**, 555 (1964).



giving the pressure, density, and so on, in each shell. In order to understand what drives the pulsation, as well as what inhibits it, Christy computes the amount of mechanical work done per cycle (that is, per complete pulsation) within each mass shell. If $\oint P\,dV$ is positive for a shell, it means that the shell is feeding energy into the pulsation. If $\oint P\,dV$ is negative, the mass shell is soaking up energy—it is performing negative work as it acts to dissipate the pulsation. The accompanying figure (after R. F. Christy) shows the P–V cycles for several mass shells. Determine the shell(s) which lead to dissipation of the pulsation and those which lead to its excitation. Numbers on the P–V cycles refer to different mass shells of star, the central shell being number 1, the surface shell, number 40.

4.3.3 An inequality important in statistical physics is

$$x \geqslant \ln(1+x)$$

Prove this relation by evaluating the areas beneath the graphs of

$$P_1(V) = \frac{1}{V} \qquad \text{and} \qquad P_2(V) = 1$$

over the interval from $V = 1$ to $V = 1 + x$. Sketch the functions and the areas. (You must sketch them in order to see the inequality!)

.....

.....

4.4 MAGNETIC WORK

When a gas expands it performs $P\,dV$ work against its surroundings. This work of expansion is a familiar example of thermodynamic work. Magnetic work is a far less familiar example, but one of growing importance. Whereas a gas can be cooled by being allowed to expand and perform work at the expense of its internal energy, certain types of solids can be cooled by allowing them to perform magnetic work. The technique of magnetic cooling is especially valuable at temperatures below 1 K. A detailed study of magnetic cooling is presented in Chapter 9. In this section we derive the magnetic analog of $P\,dV$, an expression showing how magnetic work can be expressed in terms of thermodynamic variables.

A long thin solenoid can be used to set up a strong and reasonably uniform magnetic field. The relationship between the field strength H and the solenoid current I follows from the integral form of Ampere's law

$$\oint \mathbf{H} \cdot \mathbf{dl} = \frac{4\pi}{c} I_{\text{enc}} \tag{4.9}$$

where I_{enc} is the total current enclosed by the loop indicated in Figure 4.7. The

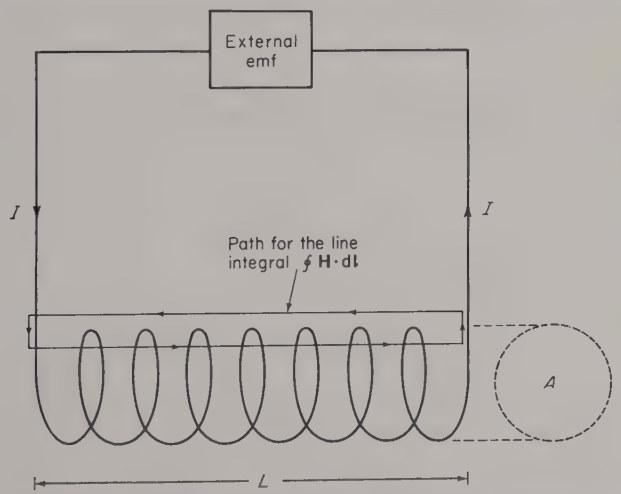


Figure 4.7. A long thin solenoid, having many turns per unit length. The leakage of the magnetic field from within the solenoid is negligible. The cross-sectional area of the solenoid is A.

total current enclosed is NI, where N is the total number of turns of the solenoid. The loop integral has the value,

$$\oint \mathbf{H} \cdot \mathbf{dl} = HL$$

where L is the length of the solenoid. This assumes that the magnetic field is negligible outside the solenoid and uniform within. The magnetic field is thus

$$H = \frac{4\pi NI}{cL} \tag{4.10}$$

If the volume V of the solenoid is filled with a magnetic material, the magnetic induction B therein is

$$B = H + \frac{4\pi M}{V} \tag{4.11}$$

where M is the total magnetic moment of the material.†

If the external emf is changed so that the current increases from I to $I+dI$, three things happen:

(a) the magnetic field changes, as per (4.10),

(b) the change in H causes M to change. If we denote these changes by dH and dM, it follows from (4.11) that the change in B is

$$dB = dH + \frac{4\pi\, dM}{V} \tag{4.11a}$$

(c) the magnetic flux threading the solenoid changes by an amount

$$d\Phi = NA\, dB \tag{4.12}$$

The changing flux induces a "back" emf (Lenz on guard!). It is this back emf which performs work. The emf is given by Faraday's law

$$\mathscr{E}_{\text{back}} = -\frac{1}{c}\frac{d\Phi}{dt}$$

The rate at which the system does work is‡

$$\frac{dW}{dt} = \mathscr{E}_{\text{back}} I = -\frac{I}{c}\frac{d\Phi}{dt}$$

---

† Here M arises through the collective response (to H) of the magnetic moments of the electrons, ions, and so on. Thermodynamics does not require that we introduce any such microscopic picture—it merely accepts M as a relevant and measurable quantity. The magnetic moment per unit volume M/V is often referred to as the magnetization.

‡ Recall that electric power is the product of voltage and current.

Using (4.12) gives

$$dW = -\frac{NIA}{c}dB \tag{4.13}$$

as the work done by the system as the induction changes from B to $B+dB$. From (4.10) we have

$$\frac{NIA}{c} = \frac{HLA}{4\pi} = \frac{HV}{4\pi}$$

where LA is the solenoid volume V, whence,

$$dW = -\frac{VH\,dB}{4\pi} \tag{4.13a}$$

This is one form for the work done *by* the system as a consequence of a change in the magnetic induction. An alternate form, which follows by using (4.11a) to replace dB by $dH+4\pi\,dM/V$, is

$$dW = -V\frac{H\,dH}{4\pi} - H\,dM \tag{4.14}$$

The first term on the right side represents the additional energy stored in the magnetic field by virtue of the increase in the solenoid current. The second term is the work done against the external source in changing the magnetic moment of the material. There is a question as to what should appear for the magnetic work done by the system in the differential form of the first law. This question is readily answered once we decide what constitutes the system. If we take the system to be the magnetic material plus the magnetic field, then (4.13a) gives the magnetic work. Often, however, the material alone is regarded as the thermodynamic system and the term $VH\,dH/4\pi$, representing energy stored in the magnetic field, is dropped. We adopt the latter viewpoint—the material alone constitutes the system. The magnetic work done *by* the system is thus

$$dW = -H\,dM \tag{4.15}$$

The magnetic analog of $P\,dV$ is therefore $-H\,dM$. The correspondences are

$$H \rightarrow P \qquad \text{(both intensive variables)}$$

$$-M \rightarrow V \qquad \text{(both extensive variables)}$$

The minus sign appearing in $-H\,dM$ arises because a positive amount of work is performed by the system when M decreases. For a magnetic system the first law of thermodynamics reads

$$dQ = dU - H\,dM + P\,dV \tag{4.16}$$

In practice the $P\,dV$ work is often ignorable compared with $-H\,dM$.

....

4.4.1 (a) Argue that the total work performed by a magnetic material during a reversible process is given by

$$W = -\int_i^f H\, dM$$

Present a geometrical interpretation of the integral.

(b) Explain why the total heat absorbed (reversibly) by a magnetic material during a cyclic process is given by

$$Q = -\oint H\, dM$$

4.4.2 The relation between M and H is generally written

$$\frac{M}{V} = \chi H$$

where χ is the magnetic susceptibility. In general χ depends on both H and T. However, for weak fields and temperatures well above the liquid helium range many paramagnetic† materials obey Curie's law,

$$\chi = \frac{C}{T}$$

where C is the Curie constant, not the heat capacity.

Consider the following reversible processes, both of which result in the doubling of M:

(i) an isothermal process in which the magnetic field strength is doubled,

(ii) a process carried out with H held constant, during which the temperature is reduced to half its original value.

Assume that the volume remains fixed during both processes.

(a) Indicate the paths of the two processes on an H–M diagram (the magnetic analog of a P–V diagram).

(b) Determine which of the two processes results in the larger performance of work by the surroundings. You may do this analytically or by an appeal to the geometrical representation of magnetic work.

....

† Paramagnetic materials have small positive values of χ and are attracted by the poles of a strong magnet. Diamagnetic materials have small negative values of χ and are repelled by a magnetic field.

4.5 HEAT CAPACITIES AND INTERNAL ENERGY

Under most circumstances the temperature of a system changes when it absorbs heat. Exceptions arise during a change of phase, as when a solid melts and becomes a liquid. The first law shows that the heat is used to change the internal energy and/or to perform work. The equations $dW = P\,dV$ and $dW = -H\,dM$ reveal how the thermodynamic work can be expressed in terms of the thermodynamic variables. In this section we investigate the relationship between internal energy changes and the other thermodynamic variables. We shall discover that the internal energy is closely related to the heat capacity.

In Chapter 3 we introduced the specific heat capacity C which relates the heat absorbed and the temperature change of a sample of mass m, namely,

$$Q = mC\Delta T \tag{3.1'}$$

C is an intensive quantity—the heat-capacity per gram. Two quantities, closely related to C, are frequently encountered. The total heat capacity is simply the mass of the system times the heat capacity per gram,

$$\text{total heat capacity of } m \text{ grams} = mC$$

When the mass m equals the molecular weight the total heat capacity is referred to as the molar heat capacity. In all future developments we shall write C for the total heat capacity instead of mC. In certain instances the molar heat capacity will also be designated simply as C. Where confusion might arise the meaning of the symbol C will be noted explicitly. With this convention in force (3.1′) would appear as $Q = C\Delta T$, or in differential form as $dQ = C\,dT$. The total heat absorbed or rejected in a finite change of temperature is

$$Q = \int dQ = \int_{T_i}^{T_f} C\,dT \tag{4.17}$$

In general, the heat capacity depends on T. If the heat capacity does not change significantly over the temperature range T_i to T_f, C may be treated as a constant in the integration which then gives

$$Q = C(T_f - T_i) = C\Delta T \tag{4.18}$$

Happily, many substances have nearly constant heat capacities over sizeable temperature ranges, permitting use of the elementary relation $Q = C\Delta T$ in lieu of the more involved integral relation (4.17).

....

4.5.1 (a) Determine the amount of heat required to raise the temperature of 50 gm of aluminum from 25 °C to 35 °C. Assume the heat capacity is

TABLE 4.1

Specific Heat Capacity at Constant Pressure[a]

Substance	C_p (cal/gm-°C)	Substance	C_p (cal/gm-°C)
Silver (Ag)	0.056	Argon (Ar)	0.125
Aluminum (Al)	0.215	Benzene (C_6H_6)	0.388
Gold (Au)	0.031	Bromine (Br)	0.107
Bismuth (Bi)	1.029	Carbon dioxide (CO_2)	0.229
Copper (Cu)	0.092	Chlorine (Cl_2)	0.115
Iron (Fe)	0.107	Ethyl alcohol (C_2H_5OH)	0.580
Mercury (Hg)	0.033	Helium (He)	1.252
Potassium (K)	0.184	Hydrogen (H_2)	3.393
Magnesium (Mg)	0.248	Methyl alcohol (CH_3OH)	0.603
Sodium (Na)	0.295	Neon (Ne)	0.246
Nickel (Ni)	0.103	Nitrogen (N_2)	0.249
Lead (Pb)	0.031	Oxygen (O_2)	0.219
Zinc (Zn)	0.084	Silicon (Si)	0.173

[a] Values refer to a pressure of 1 atm and a temperature of 25 °C for solids and liquids; 15 °C for gases. Adapted with permission from "International Critical Tables of Numerical Data: Physics, Chemistry, and Technology." Nat. Res. Council, Washington, D.C., 1923–1933, and E. H. Kennard, "Kinetic Theory of Gases." McGraw-Hill, New York (1938).

constant. Consult Table 4.1 for data. Bear in mind that Table 4.1 lists specific heats.

(b) If the amount of heat computed in part (a) were added to 50 gm of lead what would be its temperature change?

4.5.2 At sufficiently low temperatures the heat capacities of many crystalline solids are proportional to T^3. This empirical fact finds theoretical justification in the Debye theory of solids which predicts a molar heat capacity given by

$$C = 464\left(\frac{T}{\theta}\right)^3 \quad \frac{\text{cal}}{\text{mole-K}}$$

where θ is known as the *Debye temperature* and is a characteristic of the material. For temperatures low enough to permit use of the Debye form of C show that the heat absorbed in raising the temperature of one mole from T_i to T_f is

$$Q = 116\theta\left[\left(\frac{T_f}{\theta}\right)^4 - \left(\frac{T_i}{\theta}\right)^4\right]$$

....

In most laboratory situations heat exchange takes place at constant pressure. On the other hand, kinetic theory and statistical mechanics generate expressions for the heat capacity for a system of constant volume. Still other variants of the heat capacity are encountered in dealing with magnetic systems. We will use subscripts to designate which thermodynamic variable remains constant during the heat exchange. Thus

$$(dQ)_V = C_v\,dT \tag{4.19}$$

indicates a constant volume exchange, C_v being the heat capacity at constant volume. Similarly,

$$(dQ)_P = C_p\,dT \tag{4.20}$$

signifies a heat transfer at constant pressure. The specific heat capacities at constant pressure for a number of substances are displayed in Table 4.1.

Consider a simple system of fixed mass for which P, V, and T serve as the thermodynamic variables. These three variables are related by an equation of state so that only two are independent. Let V and T be taken as the independent pair. The pressure P is thereby relegated to the status of a dependent variable. The internal energy is a state function and therefore can be expressed in terms of V and T. The change in U, resulting from changes dT and dV, is given by

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV \tag{4.21}$$

Inserting this form of dU into the first law, $dQ = dU + P\,dV$, gives

$$dQ = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[P + \left(\frac{\partial U}{\partial V}\right)_T\right] dV \tag{4.22}$$

First, consider a reversible process in which heat is added while holding the volume constant. On the left side we can set

$$dQ = (dQ)_V = C_v\,dT$$

On the right side we have $dV = 0$, that is, there is no change of volume. This leaves

$$C_v\,dT = \left(\frac{\partial U}{\partial T}\right)_V dT$$

whereupon we conclude that the heat capacity at constant volume is related to the internal energy by

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V \tag{4.23}$$

The manipulations used to obtain this relation will be employed again. It is recommended that the student become familiar with the procedures.

Next, consider the addition of heat at constant pressure. The first law may be written

$$dQ_P = C_p dT = C_v dT + \left[P + \left(\frac{\partial U}{\partial V}\right)_T\right] dV \tag{4.24}$$

Since we must determine what happens at constant pressure we change our minds and regard T and P as the independent variables. The volume then becomes a variable dependent on P and T and we can write

$$dV = \left(\frac{\partial V}{\partial T}\right)_P dT + \left(\frac{\partial V}{\partial P}\right)_T dP$$

At constant pressure, $dP = 0$ and

$$(dV)_P = \left(\frac{\partial V}{\partial T}\right)_P dT$$

Inserting this into (4.24) produces the result

$$C_p = C_v + \left[P + \left(\frac{\partial U}{\partial V}\right)_T\right]\left(\frac{\partial V}{\partial T}\right)_P \tag{4.25}$$

Remarkably enough, the tangle of terms appearing beyond C_v in (4.25) can be expressed rather simply in terms of directly measurable properties of the system. The relationship between C_p and C_v is especially simple for an ideal gas. The ideal gas is "defined" by two equations:

(a) the equation of state

$$PV = nRT$$

(b) the condition that the internal energy is independent of the volume†

$$\left(\frac{\partial U}{\partial V}\right)_T = 0 \tag{4.26}$$

The second condition reduces (4.25) to

$$C_p = C_v + P\left(\frac{\partial V}{\partial T}\right)_P$$

† The experimental basis for this assumption is provided by the classic free expansion experiments of Gay-Lussac and Joule and Kelvin. These are discussed in Chapter 6.

From the equation of state we find, by differentiating with P held constant,

$$P\left(\frac{\partial V}{\partial T}\right)_P = nR$$

If we take $n = 1$ mole, there results

$$C_p = C_v + R \tag{4.27}$$

in which C_p and C_v now denote molar heat capacities.

For an ideal gas, (4.23) and (4.26) combine with (4.21) to give

$$dU = C_v\,dT \tag{4.28}$$

Experimentally it is found that gases which exhibit the "ideal" characteristics defined by (a) and (b) also have nearly constant values of C_v. With C_v constant, (4.28) can be integrated to give the internal energy

$$U = C_v T \tag{4.29}$$

In all future encounters with the ideal gas we take $U = C_v T$ for the internal energy.

The fact that C_p exceeds C_v is readily understood. Note that C_v and C_p are numerically equal to the amount of heat required to raise the temperature of 1 mole of gas by 1 °C at constant volume and constant pressure, respectively. When the gas is heated at constant volume all of the heat is used to raise the temperature of the gas. The gas does no work when maintained at constant volume $dW = P\,dV = 0$. If the same change of temperature is to be accomplished at constant pressure, more heat is required because the gas expands and does work. It was the fact that C_p exceeds C_v which led Robert Mayer to the realization that heat, like work, is but another of the many forms assumed by energy. Mayer saw that heat was being converted into work. Such a transformation implies that heat and work are merely different aspects of the same entity—energy.

Example 4.2 Adiabatic Transformations of an Ideal Gas

An adiabatic process is one in which there is no heat exchange between the system and its surroundings. We now show that the path followed by an ideal gas during an adiabatic expansion or compression is determined by the equation

$$TV^{\gamma-1} = \text{constant}$$

If we set $dQ = 0$, the first law reads

$$dU + dW = 0 \tag{4.30}$$

For an ideal gas

$$dU = C_v dT \tag{4.28}$$

For dW we write $P\,dV$ and use the equation of state to replace P by RT/V. Inserting these results into (4.30) and dividing by $C_v T$ leaves

$$\frac{dT}{T} + \frac{R}{C_v}\frac{dV}{V} = 0 \tag{4.31}$$

We recall the results

$$d(\ln x^{\alpha}) = \alpha \frac{dx}{x}$$

and

$$\ln AB = \ln A + \ln B$$

to recognize (4.31) as

$$d(\ln[TV^{\gamma-1}]) = 0 \tag{4.32}$$

with

$$\gamma \equiv 1 + \frac{R}{C_v} \tag{4.33}$$

It follows that $\ln(TV^{\gamma-1})$ and thus $TV^{\gamma-1}$ itself remain constant during the adiabatic expansion or compression. This relationship between T and V describes the path of the process. We can write

$$TV^{\gamma-1} = T_0 V_0^{\gamma-1} \tag{4.34}$$

where T_0 and V_0 are the temperature and volume of the gas for some particular state. For example, if T_0 and V_0 denote the initial temperature and volume, (4.34) relates the temperature and volume for all other states reached by an adiabatic expansion or compression. By using the equation of state, the path can be described in terms of any two of the three variables P, V, and T (see Problem 4.5.3). ✧

....

4.5.3 Using the ideal gas equation of state, show that an adiabatic path can also be described by

$$PV^{\gamma} = P_0 V_0^{\gamma}$$

and by

$$P^{1-\gamma}T^{\gamma} = P_0^{1-\gamma}T_0^{\gamma}$$

4.5.4 Wreckage of a submarine was photographed at a great depth. The wreckage showed distinct signs of charring and other indications of fire. If plates on the vessel failed, the inrushing water would adiabatically compress the trapped air. If the initial temperature was 300 K, determine the temperature when the volume of the air decreased to $\frac{1}{8}$ its initial value. Compare the temperature with the melting point of lead (600 K). Determine the pressure in atmospheres, assuming an initial pressure of 1 atm. Take $\gamma = \frac{7}{5}$, a value appropriate for an ideal diatomic gas.

4.5.5 An ideal gas expands, doubling its volume. Among the many ways this can be accomplished are: (i) isobaric; (ii) isothermal; (iii) adiabatic. Determine which of these types of expansion:

(a) results in the largest and smallest changes in the gas temperature;
(b) results in the gas performing the largest and smallest amounts of work against the surroundings;
(c) requires the largest and smallest heat absorption by the gas.

Where possible, use qualitative arguments based on the P–V diagram, knowledge of $U(T)$, and so on, rather than quantitative calculations.

....

Example 4.3 Dalton's Law of Partial Pressures

In a mixture of gases, each molecular species contributes to the total pressure. Dalton's law states that if the gases are sufficiently dilute (so that each behaves as an ideal gas), the total pressure is the sum of the *partial pressures*. The partial pressure of each component species is defined as the pressure which it would exert if it alone filled the container. If P_i denotes the partial pressure of the ith component, we can state Dalton's law of partial pressures as (K is the number of component gases)

$$P = \sum_{i=1}^{K} P_i \tag{4.35}$$

Dalton's law is valid only to the extent that each component of the mixture behaves as an ideal gas. If there are n_i moles of the ith component, its partial pressure is given by

$$P_i V = n_i RT$$

It then follows from the statement of Dalton's law (4.35) that the mixture obeys the ideal gas equation of state, that is,

$$PV = \left(\sum_i P_i\right) V = \left(\sum_i n_i\right) RT = nRT$$

where n is the total number of moles of gas in the mixture

$$n = \sum_i n_i$$

Thus, so long as each component behaves as an ideal gas the mixture itself exhibits an ideal gas character. ✧

....

4.5.6 Show that the partial pressure of the ith component of an ideal gas mixture may be expressed as

$$P_i = x_i P$$

where x_i is the mole fraction of the ith component, $x_i \equiv n_i/n$.

....

REFERENCES

The concept of reversibility is explored in:

H. C. Van Ness, "Understanding Thermodynamics," Chapter 3. McGraw-Hill, New York (1969).

A careful and exhaustive development of the concept of thermodynamic work is given by:

M. W. Zemansky, "Heat and Thermodynamics," 5th ed., Chapter 3. McGraw-Hill, New York (1968).

The thermodynamics of magnetic systems is notorious for its ambiguous treatments. The following note, which makes special reference to superconductors, provides a convenient comparison and summary of thermodynamic relations in a magnetic field:

J. L. Mundy and V. L. Newhouse, Thermodynamic Systems Involving Magnetic Fields, *Am. J. Phys.* **34**, 1195 (1966).

A broad survey of the theoretical and experimental aspects of specific heats is given in:

E. S. R. Gopal, "Specific Heats and Low Temperatures." Plenum, New York (1966).

Despite the title, the treatise is not limited to the low-temperature regime. A treatment of magnetic specific heats is included.

Chapter 5

The Second Law of Thermodynamics

5.1 INTRODUCTION

The first law of thermodynamics places a quantitative restriction on energy-conversion processes. Such transformations can alter the form but not the amount of energy. The second law of thermodynamics imposes limits on processes which convert heat into work. The conversion of heat into work is a matter of vital importance to all phases of technology. The upshot of the second law is that the complete conversion of heat into work is not possible.

The discussions which follow make frequent reference to heat engines. A *heat engine* is any device which, operating in a cycle, absorbs heat, converts part of it into work, and rejects the remainder. Phrased in terms of the limitations of heat engines, the second law states that there can be no heat engine which can absorb a given amount of heat and convert it completely into work. This is the *practical* thrust of the second law. A more precise statement of the second law is given in the following section where we present the formulations of Lord Kelvin and Rudolph Clausius.

The concept of a heat reservoir must be assimilated as a preliminary step. A *heat reservoir* is a body of uniform temperature throughout, whose mass is sufficiently large that its temperature is unchanged by the absorption or rejection of heat. If an ice cube is tossed into the Atlantic Ocean, it produces no observable change in the temperature of the ocean. The Atlantic Ocean

thereby qualifies as a heat reservoir. If the same ice cube were dropped into a cup of hot tea, a noticeable temperature change would occur. If a snowflake were to fall into the tea, no temperature change would be evident. The cup of tea therefore qualifies as a heat reservoir in the latter case, but not in the former.

5.2 THE KELVIN–CLAUSIUS FORMULATIONS OF THE SECOND LAW OF THERMODYNAMICS

Precise statements of the second law were formulated by William Thomson (later Lord Kelvin) and Rudolph Clausius in the early 1850's. Although worded differently, their statements are equivalent.

The Kelvin statement of the second law is this:

> It is impossible to devise a process whose *only* result is to convert heat, extracted from a *single* reservoir, entirely into work.

Evidently the word "only" in Kelvin's statement is an important qualifier. Many processes can be performed whereby a system converts heats completely into work. However, all such processes result in a final state which differs from the initial state. There is some *other* result beyond the conversion of heat into work. "Single" is also a key word. Heat engines convert heat into work, that is, into other forms of energy. A heat engine extracts heat from high temperature reservoirs, converts *part* of it into work, and rejects the remainder to lower-temperature reservoirs. Converting the net heat absorbed into work does not violate the second law, and is in fact demanded by the first law. A violation of Kelvin's statement of the second law would result if a heat engine could extract heat from a single reservoir and convert it *entirely* into work—without the rejection of a net amount of heat to other reservoirs.

The Clausius statement of the second law is this:

> It is impossible to devise a process whose *only* result is to extract heat from a reservoir and transfer it to a reservoir at a higher temperature.

The equivalence of the two statements may be stated somewhat cryptically as

> If Kelvin false, then Clausius false,
> and
> If Clausius false, then Kelvin false.

Of course, demonstrating their logical equivalence does not prove either version of the second law. The proof is very definitely "in the eating of the

pudding." Along with other laws of nature, the second law of thermodynamics survives only because it is in accord with experiment. No amount of experimental evidence could ever establish it as true. The evidence supporting the second law is the failure of all attempts to construct perpetual motion machines of the second kind.† A few examples best illustrate why fame, fortune, and scientific immortality would descend instantly upon anyone who managed to violate the second law.

The crust of our Earth, its atmosphere, and its oceans are giant heat reservoirs. Imagine an engine which could extract heat from the atmosphere and convert it into mechanical energy. The mechanical energy could be used to operate machinery and to generate electric currents. Air taken in and cooled by the extraction of heat could be used as a refrigerant. Air-conditioning systems would be a by-product of such an engine! The supply of energy in the atmosphere is virtually limitless, its content being maintained by the Sun. It would be unnecessary to mine coal or drill for gas and oil—*if* there were such engines. However, there are none! The oceans of the Earth likewise contain enormous quantities of energy. Imagine a ship with a power plant which could extract heat from sea water. The heat could be converted into work and used to propel the ship. Such wondrous heat engines are destined to remain figments of the imagination.‡

....

5.2.1 One mole of an ideal gas expands isothermally.

(a) If the gas doubles its volume, show that the work of expansion is

$$W = RT \ln 2.$$

(b) Since the internal energy of an ideal gas depends solely on its temperature, there is no change in U during the expansion. It follows from the first law that the heat absorbed by the gas during the expansion is converted completely into work. Why does this not violate the second law?

....

If the complete conversion of heat into mechanical energy is not possible, the obvious question is, "What is the optimum design for a heat engine?" The answer to this question was supplied by Sadi Carnot, a French engineer. Carnot's brilliant analysis of the problem gained for him credit for discovery

---

† Recall that a perpetual motion machine of the first kind is one which would manufacture energy—it would violate the first law of thermodynamics.

‡ The interior of the Earth is hotter than its surface. The interior is a bona fide high-temperature reservoir and is now being tapped at several geothermal power plants. See J. Barnea, Geothermal Power, *Sci. Am.* **226**, 70 (January 1972).

of the second law of thermodynamics. Carnot's efforts are especially noteworthy because his views were set forth in 1824, nearly 20 years prior to the Mayer–Joule formulation of the first law. Carnot died 8 years later in a cholera epidemic, at the age of 36. Because Carnot employed the caloric theory of heat his applications of the second law were not always sound. Carnot's notebooks, unpublished until long after his death, revealed his suspicion of the caloric theory and his resolve to perform experiments similar to those carried out by Joule.

There are two remarkable aspects of Carnot's researches. First, he found that the efficiency with which heat can be converted into other forms of energy depends on the nature of the cyclic process employed but not at all on the working substance—the material which undergoes the cyclic transformations. Second, Carnot discovered the most efficient cyclic process for converting heat into other forms of energy. This cycle, now known as the Carnot cycle, is discussed in Section 5.3. Carnot's achievements are truly monumental. Without benefit of a correct theory of heat he set forth the ultimate standard of heat engine efficiency.

There are of course many types of heat engines in use today. Basically, a heat engine operates by (a) extracting heat from a high-temperature reservoir, (b) converting part of the heat into work, and (c) rejecting the remainder to a low-temperature reservoir. It was Carnot's investigation of the efficiency of heat engines which led him to an understanding of the second law. A brief look at the Carnot cycle paves the way for a demonstration of the equivalence of the Kelvin and Clausius formulations of the second law.

....

5.2.2 It was emphasized in Section 4.5 that specific heat capacities are never negative. Adding heat to a system never causes its temperature to decrease. This is an experimental fact. Construct a scheme which would violate the second law of thermodynamics by assuming you have a material with a negative heat capacity.

5.2.3 Capillary action causes liquids to rise against the force of gravity. Suppose you have a tube in which water will rise to a height of 2 ft. Next, suppose the tubing is bent so as to form an inverted "j" 1-ft-high. Capillary action therefore has an easy time filling the entire tube. Next, one places a small paddle wheel beneath the end of the tube. As water drips onto the paddle wheel it turns and generates mechanical energy. Water which strikes the paddle wheel falls back into the reservoir with less kinetic energy than it would have had it not struck the wheel. Some of its kinetic energy has been converted into the mechanical energy of the rotating paddle wheel. This would result in a cooling of the water were it not for the fact that the reservoir is maintained

at a constant temperature by virtue of thermal contact with the Earth. So—the only result of the device is to extract heat from a reservoir (the Earth) and convert it into work. This violates the second law and so there is no point in trying to build one. Explain why the device does not operate as just advertised. [*Hint:* In case you have not already discovered the flaw it will be helpful to know that the capillary tube will not even function as a perpetual motion machine of the third kind. (A perpetual motion machine of the third kind is a device which "just runs." A frictionless wheel which spins forever once started or a pendulum which swings without damping typify perpetual motion machines of the third kind.)]

•••••• ••••••

5.3 THE CARNOT CYCLE

The Carnot cycle is a series of reversible processes. It consists of two adiabatic legs connected by a pair of isothermal legs. Figure 5.1 illustrates a Carnot cycle in which a gas serves as the working substance. The arrows indicate the sequence of steps. During the isothermal expansion from 1 to 2,

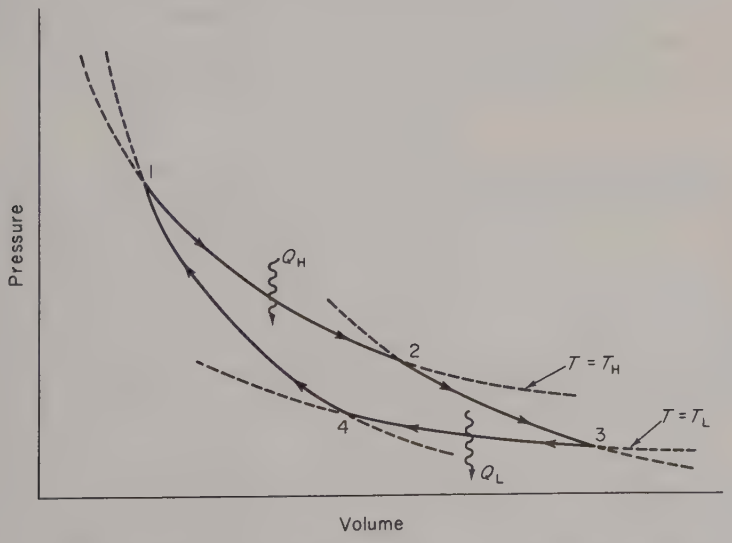


Figure 5.1. The Carnot cycle for a gas as it appears on a P–V diagram, where T_H and T_L refer to the high and low temperature reservoirs. Heat Q_H is absorbed from a reservoir as the gas expands isothermally from 1 to 2. The expansion from 2 to 3 is performed adiabatically. Heat Q_L is rejected during the isothermal compression from 3 to 4. The cycle is completed by an adiabatic compression from 4 to 1.

heat Q_H is absorbed from a high-temperature reservoir. The gas expands performing work against the surroundings. The expansion of the gas continues from 2 to 3 under adiabatic conditions, that is, the gas is insulated from its surroundings during the expansion. An isothermal compression from 3 to 4 results in the performance of positive work by the surroundings and the rejection of heat Q_L by the gas to the low-temperature reservoir. The cycle is completed from 4 to 1 by adiabatic compression. There is no heat transfer and the positive work performed by the surroundings increases the internal energy of the gas. The increase shows up as a temperature rise from T_L to T_H.

The overall result of the cycle is the conversion into work of some of the heat extracted from the high-temperature reservoir. To avoid possible misunderstandings we point out that our sign convention, for heat absorbed by the system and by the surroundings, has been temporarily abandoned, with Q_H and Q_L denoting positive quantities of heat. The rejection of heat by the system will be indicated explicitly by a minus sign. For the Carnot cycle just described the system absorbs heat Q_H and rejects an amount Q_L. The net amount of heat received from the surroundings is $Q_H - Q_L$. By the same token, the net work W done by the system over the cycle is composed of two positive contributions and two negative contributions. During the expansions, 1 to 2 and 2 to 3, the system performs positive work. During the compressions, 3 to 4 and 4 to 1, the system does negative work, that is, the surroundings perform positive work. Since the system returns to its initial thermodynamic state there is no net change in its internal energy; $\Delta U = 0$ for a cyclic process. The first law

$$Q = \Delta U + W$$

when applied to the Carnot cycle becomes

$$Q_H - Q_L = W$$

This merely states that the net work done by the system over the cycle equals the net heat intake.

Observe that only a portion of the heat absorbed at the higher temperature is converted into work. The second law of thermodynamics would crumble if it were possible to make Q_L zero.

....

5.3.1 In Figure 5.1, the adiabat legs are steeper than the isotherm segments of the Carnot cycle. This may be verified provided the equation of state and internal energy of the gas are known functions of P, V, and T. For an ideal gas (1 mole)

$$PV = RT; \qquad U = C_v T$$

where C_v is the molar heat capacity at constant volume and is a constant.

The slope dP/dV of the ideal gas adiabat is larger by a factor of $\gamma = 1 + R/C_v$ than the slope of an isotherm. Verify this by showing that

$$\left(\frac{dP}{dV}\right)_{\text{isot}} = -\frac{P}{V}; \qquad \left(\frac{dP}{dV}\right)_{\text{adia}} = -\gamma\frac{P}{V}$$

where isot is isotherm and adia is adiabat. Recall that an adiabat is defined by the condition that $dQ = 0$ and that in general $dQ = dU + P\,dV$ for a gas.

5.3.2 Suppose that an ideal gas is used as the working substance in the Carnot cycle depicted in Figure 5.1. The internal energy of an ideal gas depends on the temperature alone $U = U(T)$.

(a) If the gas absorbs $Q_{\text{H}} = 3$ kcal during the isothermal leg from 1 to 2, how much work is performed by the gas?

(b) If the gas performs 1 kcal of work during the adiabatic expansion from 2 to 3, what is the change in its internal energy?

(c) How much work must be performed on the gas from 3 to 4 for it to reject $Q_{\text{L}} = 2$ kcal of heat?

(d) How much work must be performed on the gas to take it from 4 to 1?

(e) What is the net amount of work performed by the gas during the cycle?

....

5.4 THERMODYNAMIC EFFICIENCY

We would say that a process is efficient if it converts a sizable fraction of the heat absorbed into useful work, rejecting only a small fraction. The *thermodynamic efficiency* η is defined as the fraction of the total heat absorbed which is converted into work (over a cycle)

$$\eta = \frac{W}{Q_{\text{abs}}} \tag{5.1}$$

An alternate expression for η follows from the fact that the work performed during a cyclic process equals the net heat exchange: Thus $W = Q_{\text{abs}} - Q_{\text{rej}}$, so that

$$\eta = \frac{Q_{\text{abs}} - Q_{\text{rej}}}{Q_{\text{abs}}} = 1 - \frac{Q_{\text{rej}}}{Q_{\text{abs}}} \tag{5.2}$$

For the Carnot cycle, Q_{H} is the total heat absorbed and Q_{L} is the heat rejected. The Carnot efficiency is therefore

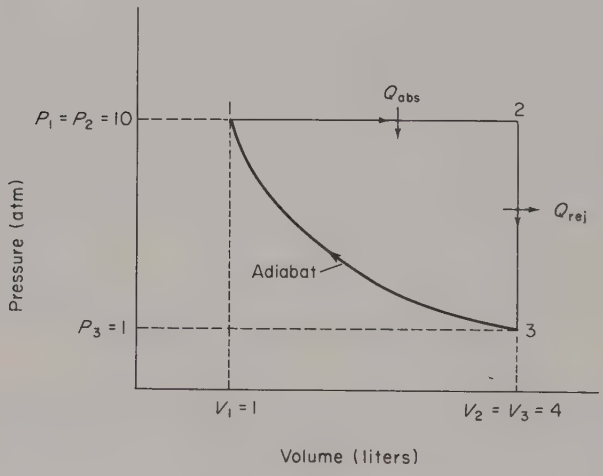
$$\eta_{\text{c}} = 1 - \frac{Q_{\text{L}}}{Q_{\text{H}}} \tag{5.2a}$$

In the next section we show that no heat engine operating between two reservoirs can be more efficient than a Carnot engine operating between the same two reservoirs.

Example 5.1 Calculation of Thermodynamic Efficiency

As an elementary illustration showing how thermodynamic efficiencies are determined we consider the three-legged cycle shown in the accompanying figure. The working substance is taken to be an ideal gas. The isobaric expansion of the gas from 1 to 2 leads to an increase in temperature and the absorption of heat

$$Q_{\text{abs}} = C_p \Delta T = C_p(T_2 - T_1)$$



The isovolumic decompression from 2 to 3 results in a temperature drop and the rejection of heat

$$Q_{\text{rej}} = -C_v \Delta T = C_v(T_2 - T_3)$$

The cycle is completed along the adiabat from 3 to 1, there being no heat exchange. From (5.2)

$$\eta = 1 - \frac{Q_{\text{rej}}}{Q_{\text{abs}}} = 1 - \frac{C_v(T_2 - T_3)}{C_p(T_2 - T_1)}$$

Recalling that $C_p/C_v = \gamma$ we have

$$\eta = 1 - \frac{1}{\gamma} \frac{\left(1 - \dfrac{T_3}{T_2}\right)}{\left(1 - \dfrac{T_1}{T_2}\right)} \tag{5.3}$$

The temperature ratios may be expressed in terms of the volume and pressure ratios by using the equation of state. As η is an intensive property of a system we consider 1 mole of gas. It follows from $PV = RT$ that an alternate expression for the efficiency is

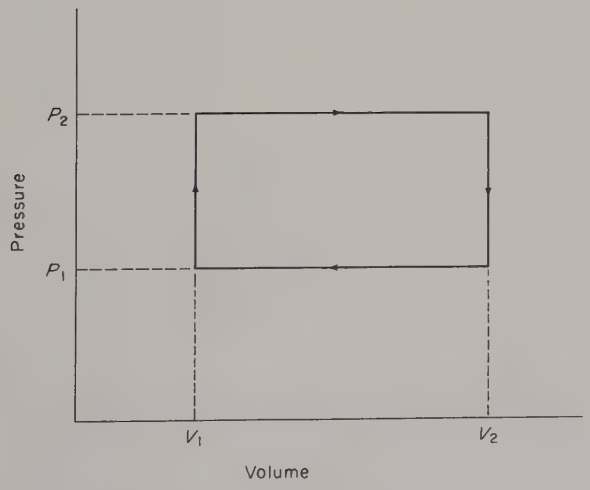
$$\eta = 1 - \frac{1}{\gamma} \frac{\left(1 - \frac{P_3}{P_2}\right)}{\left(1 - \frac{V_1}{V_2}\right)} \quad \diamondsuit \tag{5.4}$$

5.4.1 (a) Use the ideal gas equation of state to show that (5.4) follows from (5.3).

(b) Determine the efficiency of the cycle of Example 5.1, and the P–V data therein, for the cases of ideal monatomic and diatomic gases ($\gamma_{\text{mon}} = 5/3$; $\gamma_{\text{diat}} = 7/5$).

5.4.2 Show that the efficiency for the cycle shown in the accompanying figure is

$$\eta = \frac{\gamma - 1}{\frac{\gamma P_2}{P_2 - P_1} + \frac{V_1}{V_2 - V_1}}$$



It was emphasized that the Carnot cycle consists of a sequence of reversible processes. As described earlier, the Carnot cycle functions as a heat engine; it converts heat into work. If the sequence of steps is reversed, the Carnot cycle functions as a refrigerator. The working substance of a refrigerator extracts heat from a low-temperature reservoir, has (positive) work performed on it by the surroundings, and rejects heat at a higher temperature. It therefore pumps heat uphill. The efficiency of a refrigerator is defined as the fraction of the heat rejected which was received as work

$$\eta = \frac{W}{Q_{\text{abs}}} = 1 - \frac{Q_{\text{rej}}}{Q_{\text{abs}}}$$

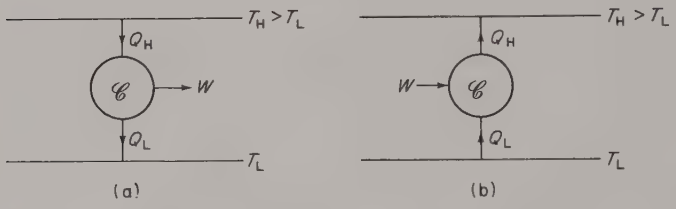


Figure 5.2. The Carnot engine (a) and refrigerator (b). An arrow directed toward the engine indicates heat absorption *by* or work performed *on* the engine.

This definition guarantees that a reversible cycle has the same efficiency whether it operates as a heat engine or as a refrigerator. Figures 5.2a and b tell the Carnot cycle story forward and backward. The definition of the thermodynamic efficiency of a refrigerator is inconvenient in practice. A more useful measure of the capabilities of a refrigerator is the coefficient of performance defined as

$$\zeta \equiv \frac{\text{heat absorbed}}{\text{work performed}} = \frac{Q_{\text{abs}}}{W}$$

Optimally, a refrigerator extracts a sizable amount of heat from a low-temperature reservoir with the expenditure of a minimal amount of work, that is, it has a large coefficient of performance. The coefficient of performance is inversely proportional to the basic economic variable of refrigeration—the cost per calorie extracted.

We discuss the Carnot cycle further in the next section. We conclude this section by demonstrating the equivalence of the Kelvin and Clausius formulations of the second law. We must show that if Kelvin's statement is false, so then is Clausius', and vice versa.

Suppose first that Kelvin's statement is false, that is, suppose that a process

is possible whose only result is to convert heat extracted from a reservoir completely into work. The work so obtained could then be completely converted back into heat at any desired temperature, via friction, for example. In particular, it could be deposited in a heat reservoir whose temperature exceeds that of the reservoir from which the heat was extracted. The only result of the scheme is to transfer a quantity of heat from a lower to a higher temperature, which is a violation of Clausius' postulate. Thus, if Kelvin is false, so is Clausius.

To complete the proof of their logical equivalence we must also show that if Clausius' version of the second law is false, so also is Kelvin's statement. Let us therefore assume that Clausius' statement is false and that there exists a scheme whose only result is to extract an amount of heat Q and transfer it to a reservoir at a higher temperature. We set up a Carnot engine which operates between the two reservoirs. Let the engine extract heat Q from the higher-temperature reservoir, convert part of it into work, and reject the remainder to the low-temperature reservoir. The size of the engine is designed such that it executes an integral number of cycles during the process. The engine therefore undergoes no change. The high-temperature reservoir likewise has undergone no change. It absorbed and subsequently rejected heat Q. If η is the efficiency of the engine, the work output is ηQ, and the heat returned to the low-temperature reservoir is $(1-\eta)Q$. Overall then the only result is that the low-temperature reservoir has given up heat ηQ, and this has been converted into work. The only result of the process is to convert into work heat extracted from a single reservoir—a violation of Kelvin's statement of the second law. This completes the proof of the equivalence of Clausius' and Kelvin's statement of the second law.

We now turn to Carnot's Theorem which, among other things, provides the basis for a thermodynamic temperature scale—a scale independent of the thermometric substance.

5.5 CARNOT'S THEOREM AND THE THERMODYNAMIC TEMPERATURE SCALE

The Carnot cycle is of special interest for two reasons. The first concerns Carnot's Theorem, which deals with the efficiency of heat engines. Carnot's Theorem proves that a heat engine which extracts heat from one reservoir, performs work, and rejects heat to a second reservoir cannot have an efficiency which exceeds that of a Carnot engine operating between the same two reservoirs. The efficiency of a Carnot engine therefore presents a theoretical limit which cannot be exceeded by clever design.

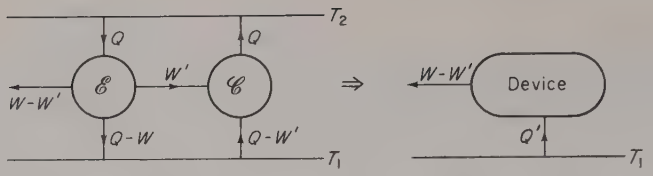


Figure 5.3. If the efficiency of $\mathscr{E}$ exceeds the Carnot efficiency, it is possible to use a portion of the work output of $\mathscr{E}$ to drive a Carnot refrigerator so as to leave the high-temperature reservoir unchanged. The only result of $\mathscr{E}$ and $\mathscr{C}$ in tandem would be to convert heat completely into work—a violation of the second law.

Carnot's Theorem is this:

> No heat engine operating between two heat reservoirs can be more efficient than a Carnot engine operating between the same two reservoirs.

The proof of the theorem follows these lines: One shows that if the theorem is false, it is possible to violate the second law. Since it is not possible to circumvent the second law, the theorem is not false. Figure 5.3 shows how the proof is formalized. We assume that the engine $\mathscr{E}$ on the left is more efficient than the Carnot engine $\mathscr{C}$ on the right. Let the engine execute an integral number of cycles in the process of extracting heat Q from the high-temperature reservoir, converting part of it into work W, and rejecting heat $Q-W$ to the low-temperature reservoir. We divide the work output into two parts, W' and $W-W'$. The work W' is used to drive the Carnot engine, operating as a refrigerator, through an integral number of cycles and such that it rejects heat Q to the high-temperature reservoir. We can always arrange to have the Carnot refrigerator go through an integral number of cycles (for a given value of Q) by adjusting its size. The efficiencies are given by

$$\eta_{\mathscr{E}} = \frac{W}{Q}, \qquad \eta_{\mathscr{C}} = \frac{W'}{Q}$$

Since, by assumption, $\eta_{\mathscr{E}} > \eta_{\mathscr{C}}$, we must have $W > W'$, and so $W-W'$ is positive, that is, the $\mathscr{E}-\mathscr{C}$ tandem does a positive amount of work against the surroundings. This series of operations leaves the high-temperature reservoir unchanged—it rejected and reabsorbed heat Q. Both $\mathscr{E}$ and $\mathscr{C}$ execute an integral number of cycles so that they are left unchanged. The low-temperature reservoir received heat $Q-W$ and rejected heat $Q-W'$ so that overall it rejected an amount of heat equal to $W-W'$. Therefore, the low-temperature reservoir rejected an amount of heat equal to the net work output of the $\mathscr{E}-\mathscr{C}$ tandem. The only result of the process is the conversion into work of heat extracted from a reservoir. However, this violates Kelvin's form of the second law. Since this is not possible we conclude that the original assumption is false—it is not possible for $\eta_{\mathscr{E}}$ to exceed $\eta_{\mathscr{C}}$. Thus, Carnot's Theorem is proved.

A corollary to Carnot's Theorem states that the efficiency of all Carnot engines (working between the same two reservoirs) is the same regardless of their working substances. To prove the corollary, merely replace the engine by a Carnot engine $\mathscr{C}_2$, whose working substance is different from that of the Carnot engine which operates as a refrigerator $\mathscr{C}_1$. Carnot's Theorem shows that

$$\eta_{\mathscr{C}_2} \text{ is not greater than } \eta_{\mathscr{C}_1},$$

that is, the "newly arrived" Carnot engine cannot have an efficiency which exceeds that of the "original" Carnot engine. Since the Carnot cycle is reversible we can turn everything around and let $\mathscr{C}_1$ operate as a heat engine which drives $\mathscr{C}_2$ now operating as a refrigerator. Carnot's Theorem now states that

$$\eta_{\mathscr{C}_1} \text{ is not greater than } \eta_{\mathscr{C}_2},$$

We conclude that

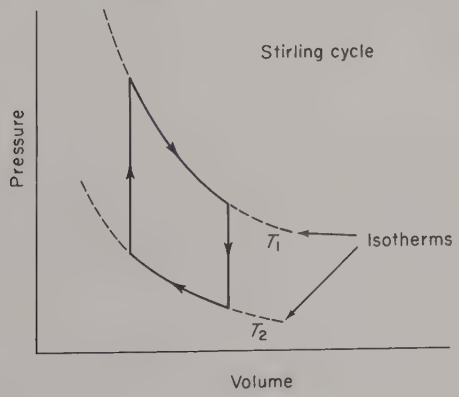
$$\eta_{\mathscr{C}_1} = \eta_{\mathscr{C}_2}$$

which proves the corollary.

Carnot's Theorem and Corollary are very important. They impose a fundamental limit on the conversion of heat into work. No amount of ingenuity can devise a heat engine more efficient than the Carnot engine. In practice it is not generally feasible to use a Carnot cycle. Among other reasons, it may be prohibitively expensive to maintain adiabatic conditions because of the elaborate insulation required. The Carnot efficiency therefore serves as a standard one aims to approach using a more economical cycle.

....

5.5.1 The Stirling engine was invented by Robert Stirling in 1816. The working substance was air. The accompanying figure shows an idealized



version of the Stirling cycle. There is renewed interest in the Stirling engine today because it holds promise as an ecologically clean automotive power plant. Note that in the Stirling cycle two isovolumic legs replace the pair of adiabatic legs of the Carnot cycle. The design of the engine is such that no heat is exchanged with the surroundings during the isovolumic steps. Instead, there is an internal exchange of heat. There are two gas-driven pistons, connected by a linkage which maintains a temperature difference while allowing the exchange of gas between piston cylinders. During the isovolumic phases of the cycle one piston moves downward, the other upward. There is an exchange of gas, and heat, between cylinders. There is no change in total volume and no heat exchange with the surroundings. With these facts in mind, verify that the efficiency of the Stirling cycle equals that of the Carnot cycle. Assume that the working substance is an ideal gas.

....

The fact that the Carnot efficiency is a universal property, independent of the working substance used in the engine, brings us to the second reason for taking another look at the Carnot cycle. In Chapter 3 we noted that the thermodynamic scale of temperature is based on the laws of thermodynamics—that it is independent of the thermometric substance used to construct a thermometer. The Carnot efficiency is independent of the working substance, but does depend on the temperatures of the heat reservoirs. This fact makes it possible to use the Carnot efficiency to define a thermodynamic temperature scale.

The efficiency of a Carnot engine which absorbs heat Q_1 along one isothermal leg and rejects Q_2 along another is

$$\eta_c = 1 - \frac{Q_2}{Q_1}$$

A thermodynamic temperature scale is defined by taking $Q_2/Q_1 = T_2/T_1$ or

$$\frac{T_2}{Q_2} = \frac{T_1}{Q_1} \tag{5.5}$$

This definition tells us that the thermodynamic temperature T is proportional to Q. Heat serves as the thermometric variable for the thermodynamic scale. This definition leaves open the size of the temperature unit. As noted in Chapter 1, the temperature of the triple point of water has been assigned the value 273.16 K. Thus if we operate a Carnot engine with one reservoir at the triple point, the temperature of the other reservoir is

$$T = 273.16 \frac{Q}{Q_3} \text{K} \tag{5.5a}$$

The thermodynamic scale is completely satisfactory from an operational viewpoint because an operational definition of heat has been provided by the first law of thermodynamics. As a consequence of (5.2) and (5.5a), the efficiency of a Carnot engine working between reservoirs at T_H and T_L is given by

$$\eta_c = 1 - \frac{T_L}{T_H} \tag{5.6}$$

In the next section we use (5.5), applied to a special Carnot cycle, to derive the Clausius–Clapeyron equation, which is extremely useful in studies of phase equilibrium.

5.6 THE CLAUSIUS–CLAPEYRON EQUATION

Our atmosphere contains water in the gaseous phase. The contribution of the water molecules to the total atmospheric pressure is called the "water vapor[†] pressure." For any given temperature there is a maximum water vapor pressure known as the "saturated water vapor pressure." The concept of a vapor pressure is not restricted to water nor to our atmosphere. The region above any liquid will contain its vapor along with, perhaps, other gases. The vapor pressure is that portion of the total pressure contributed by the vapor. There is a maximum vapor pressure, the saturated vapor pressure, whose value depends on the temperature and the physical nature of the molecules, but not on the presence or absence of other gases. The saturated vapor pressure depends solely on the temperature in a way dictated by the Clausius–Clapeyron equation. A special miniaturized Carnot cycle will be used to derive the Clausius–Clapeyron equation. We preface the derivation with a brief description of the liquid–vapor transition.

The vaporization of a liquid is readily understood from the microscopic viewpoint. By virtue of their kinetic energy some molecules are able to overcome the attractive intermolecular forces, escape the liquid, and become part of the vapor atmosphere above the liquid surface. Of course, molecules continually fall back into the liquid. Equilibrium exists—we say the vapor is *saturated*—when the rates of escape and return are equal.

The vapor exerts a pressure on the liquid. Macroscopically, this pressure arises because the liquid surface must support the weight of the vapor "piled" on it. Microscopically, the pressure results because the surface experiences an impulsive force each time a molecule falls back into the liquid. A similar force arises when a molecule "leaps" upward into the vapor atmosphere.

[†] The words "vapor" and "gas" are completely synonymous, although "vapor" is frequently used in referring to the gaseous phase in equilibrium with its liquid or solid phase.

The vapor pressure increases with temperature. The temperature at which the vapor pressure equals 1 atm (760 Torr) is termed the *normal boiling point*, and the temperature at which the vapor pressure equals the ambient pressure above the liquid is called the *boiling point*. By reducing the pressure above the liquid (climb a mountain, reduce pressure via pump, and so on), the boiling point may be lowered. For example, the normal boiling point of liquid helium is 4.2 K. If the pressure of the liquid is reduced by pumping away the vapor rapidly, the temperature of the liquid is reduced, the thermal energy being carried away with the vapor as *latent heat* of vaporization.† Temperatures down to about 0.8 K may be produced in this fashion. Below 0.8 K, an unusual film creep phenomenon peculiar to ^{4}He prevents further cooling. The rare isotope ^{3}He may be cooled by forced evaporation to temperatures near 0.3 K.

•••••• ••••••

5.6.1 How does the presence of water vapor affect the air pressure? For equal air temperatures, would the air pressure be greater over an arid desert-like surface, or over an ocean?

•••••• ••••••

To understand the vapor–liquid phase transition, consider a gas at a temperature T confined within a cylinder by a piston. The gas is to be compressed isothermally. This is accomplished by maintaining thermal contact between the contents of the cylinder and a heat reservoir at temperature T. The isothermal compression leads to an increase in pressure. The relationship between P and V along an isotherm of the gaseous phase is qualitatively similar to that for an ideal gas

$$PV = nRT$$

In particular, an isothermal compression increases the pressure. Provided the temperature is not too high, a real gas will reach a pressure at which condensation begins. Referring to Figure 5.4 we can follow the compression of the vapor from 1 to 2 along the isotherm labeled T. At 2 we have a saturated vapor.

---

† The latent heat of vaporization is the amount of heat required to convert a substance (at its boiling point) from a liquid to a vapor. The adjective latent was introduced by Joseph Black early in the 19th century. Ordinarily the addition of heat produces a temperature rise. This heat, Black referred to as a sensible heat. Black also discovered that the addition of heat could produce a change of phase, but that no temperature change occurred during the change of phase. To describe this, Black coined the terminology "latent heat." Latent heats are associated with vaporization, melting, and sublimation. The latent heat of fusion is the amount of heat required to convert a substance (at its melting point) from a solid to a liquid. The latent heat of sublimation is the amount of heat required to convert a substance from the solid phase directly to the gaseous phase.

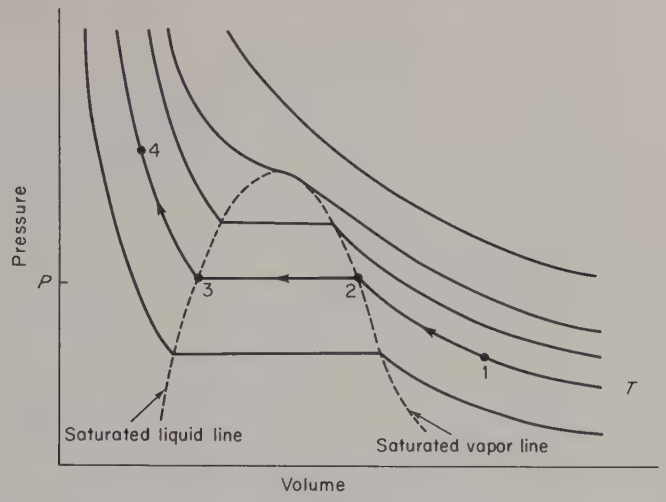


Figure 5.4. Typical isotherms in the condensation region. (See text for details.)

As the piston reduces the volume of the system, vapor continues to condense, while the pressure and temperature remain constant. The constancy of pressure and temperature is a general characteristic of condensation. The horizontal segment of the isotherm represents states in which both liquid and vapor are present. The pressure for the horizontal segment (2 to 3) of the isotherm is called the "saturated vapor pressure at a temperature T." Continued compression eventually reduces the volume to the point 3 where there is no vapor, that is, the vapor has condensed completely and the volume of the system equals the volume of the liquid. At 3 we have a saturated liquid. The isotherm rises sharply for the liquid phase, 3 to 4, because liquids are nearly incompressible.

The process of condensation is reversible. By expanding the volume of the system, the liquid may be evaporated. Bear in mind that the system must be in thermal contact with a heat reservoir which supplies or absorbs whatever heat is required to maintain a constant temperature—it is not just a matter of moving the piston to expand or compress the system.

The locus of saturated vapor points forms the saturated vapor line of Figure 5.4. Likewise, the saturated liquid line is the locus of saturated liquid states. The *critical point* marks their common terminus. At the critical point there is no distinction between liquid and vapor—their properties being identical. At temperatures above the critical temperature there is no condensation—compression leads to a continuous transition from the gaseous to liquid state. Further discussion of phase transitions is presented in Chapter 9.

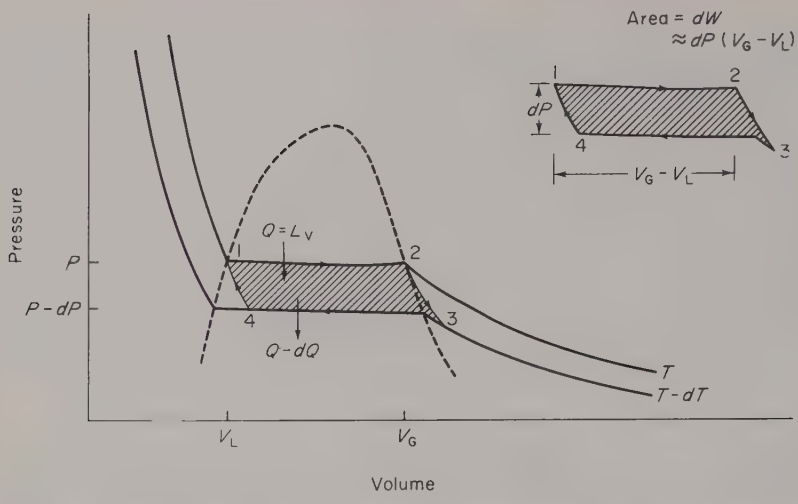


Figure 5.5. Special Carnot cycle and approximate work area. The dashed line outlines the dome portion of the P–V–T surface where liquid and vapor can coexist in equilibrium.

The Clausius–Clapeyron equation relates the vapor pressure P and temperature T, by giving the slope dP/dT of the P–V–T surface over the liquid–vapor coexistence region. We now derive the Clausius–Clapeyron equation by considering a very special Carnot cycle. Let 1 mole of the fluid be carried through the Carnot cycle indicated in Figure 5.5. The cycle is started at 1, with 1 mole of liquid (volume V_L) and no vapor present. The isothermal expansion from 1 to 2 requires the addition of heat Q. At 2, the liquid has completely evaporated so that we have 1 mole of vapor (volume V_G) and no liquid. The heat added Q is therefore equal to L_v the (molar) latent heat of vaporization. The adiabatic expansion ($2 \rightarrow 3$) lowers the temperature and pressure infinitesimally (to $P-dP$ and $T-dT$). The gas is then compressed isothermally ($3 \rightarrow 4$) and the cycle completed by an adiabatic compression to the initial state. The area on the P–V diagram enclosed by the cycle equals the work done by the system during the cycle. Approximating[†] this area by a rectangle (compare Figure 5.5) gives

$$dW = dP(V_G - V_L) \tag{5.7}$$

where V_G and V_L are the volumes occupied by 1 mole of gas and liquid, respectively. Because the system undergoes a Carnot cycle between

[†] The approximation would become exact if a more rigorous but cumbersome limiting process were employed.

temperatures T and $T-dT$ we have by (5.5),

$$\frac{Q-dQ}{Q} = \frac{T-dT}{T}$$

which gives

$$dQ = Q\frac{dT}{T}$$

However, in a cyclic process, $dQ = dW$, that is, the net heat absorbed equals the work done. Thus

$$Q\frac{dT}{T} = dP(V_G - V_L) \tag{5.8}$$

Remembering that $Q = L_v$, the latent heat of vaporization, and solving for dP/dT gives the Clausius–Clapeyron equation

$$\frac{dP}{dT} = \frac{L_v}{T(V_G - V_L)} \tag{5.9}$$

The same sort of analysis can be applied to the solid–liquid transition. The temperature variation of the melting pressure is governed by

$$\frac{dP}{dT} = \frac{L_f}{T(V_L - V_S)} \tag{5.9a}$$

where L_f is the latent heat of fusion, T is the melting point, and $V_L - V_S$ is the solid to liquid change of volume.

....

5.6.2 Compute the change (including sign) in the temperature at which water freezes when the pressure is raised from 1 atm to 100 atm.

....

It is apparent from Figure 5.5 that V_G and V_L depend on temperature. Not so apparent is the fact that the latent heat L_v also varies with temperature. By invoking some reasonable approximation, or using experimental data, it is possible to integrate the Clausius–Clapeyron equation and obtain a relation between P and T. A crude but often adequate approximation is to ignore V_L in comparison with V_G and then assume that the vapor obeys the ideal gas law $PV_G = RT$. With these approximations inserted into (5.9) we obtain

$$\frac{dP}{P} = \left(\frac{L_v}{R}\right)\frac{dT}{T^2} \tag{5.10}$$

If L_v is assumed constant, integration gives

$$\ln P = -\frac{L_v}{RT} + \text{constant} \tag{5.11}$$

A more refined version of this equation serves as a useful low-temperature thermometer. The vapor pressure above liquid helium is readily measured and, since it depends on temperature alone, serves to indicate the liquid temperature.

Example 5.2 Latent Heat of Vaporization of Mercury

According to (5.11), a plot of $\ln P$ versus $1/T$ should give a straight line of slope $-L_v/R$. Thus by measuring both the vapor pressure and the temperature it is possible to determine the latent heat of vaporization. Let (P_1, T_1) and (P_2, T_2) denote corresponding pairs of data. From (5.11)

$$L_v = \frac{R \ln\left(\frac{P_2}{P_1}\right)}{\frac{1}{T_1} - \frac{1}{T_2}}$$

We use P–T data for mercury taken from the Chemical Rubber Co. Handbook,[†]

$$P_2 = 14.8, \qquad T_2 = 470\,\text{K}$$

$$P_1 = 0.1, \qquad T_1 = 354\,\text{K}$$

These data give $L_v \simeq 14.4$ kcal/mole, tolerably close to the precision experimental value of 13.99 kcal/mole. ✧

The next section presents Maxwell's demon and a distinct change of pace. Our concern still lies with the second law, but our viewpoint shifts to the microscopic, molecular level. On to the demon!

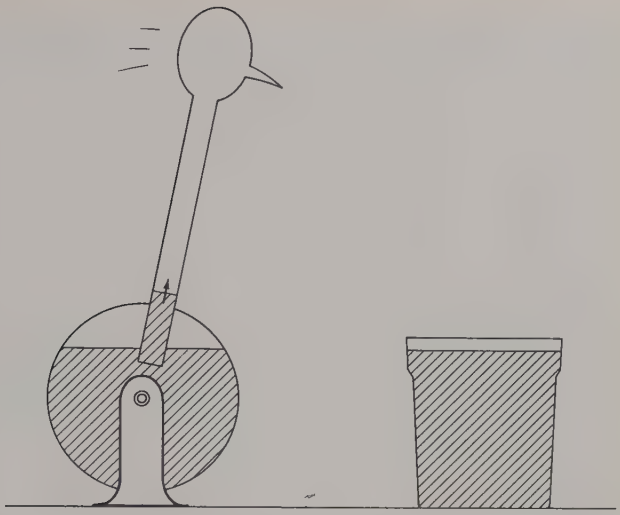
....

5.6.3 Consult tables giving vapor pressure data for water. Compute the latent heat of vaporization per mole and per gram. The experimental value at 100 °C is 539 cal/gm.

5.6.4 The "drinking duck" appears to be a perpetual motion machine. The accompanying figure shows the internal structure of the duck. The beak of the

---

[†] "Handbook of Chemistry and Physics," 51st ed. Chem. Rubber Pub. Co., Cleveland, Ohio (1970).



duck is covered with a thin layer of very absorptive fabric, much like an ink blotter. If the beak is moistened and atmospheric conditions are such as to permit the liquid to evaporate, the duck begins to sway on its perch, and finally tips forward to "drink." It then rights itself. So long as the beak is moist the drinking action continues. The fluid filling the duck is highly volatile (usually ethyl ether). Explain the principle of the duck's operation.

5.6.5 The air in a person's lungs is saturated with water vapor. What is the water vapor pressure in the lungs? (This problem will require the student to consult tables not contained in this text.)

....

5.7 MAXWELL'S DEMON

The labor saving perpetual motion machines conjured up by wishful thinking in Section 5.2 revealed the monumental consequences of any scheme which could violate the second law. A great many clever people have sought to circumvent the second law. Many of these people were crackpots with at most a very superficial understanding of physics. A most notable exception was the scheme proposed by James Clerk Maxwell, the brilliant physicist who synthesized the theory of electromagnetism. Maxwell invented a "demon." Rightfully, it is known as "Maxwell's demon." The demon would violate the second law by making heat flow from a cold body to a hot body. Maxwell's

proposal was not idle speculation. He was not trying to set up a straw man (demon!) for others to joust with. Maxwell's demon relied on the kinetic theory of gases. At the time, kinetic theory was in its infancy, but was developing fruitfully in the hands of Clausius, Maxwell, and Boltzmann. Kinetic theory operates at the microscopic level. The second law seemed firmly established at the macroscopic level, but what about the underlying microcosm? Perhaps a new wealth of science lurked deep within the breadbox-sized objects with which we deal from day to day. Perhaps it would be possible to violate the second law by judicious engineering at the microscopic level. Toward the end of his famous textbook Maxwell therefore posed the question as to whether or not it might be possible to sort the molecules of a gas in such a way as to violate the second law. Maxwell says, "But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, ...,"† and so the demon was born. We now describe Maxwell's demon and explain how its *modus operandi* supposedly violates the second law.

Two results of the kinetic theory of gases must be inserted here without proof. Their justification comes later, in Chapters 11 and 12. First, there is a *distribution* of molecular speeds in a gas, that is, not all molecules move at the same speed. Some travel at very low speeds, some at very high speeds; the majority travel at speeds comparable with the average molecular speed. Second, we need to know that the temperature of a gas is increased by any mechanism which increases the average speed of the molecules—the higher the average speed, the hotter the gas. Let us accept these two kinetic theory results and put the demon to work.

When the demon reports for work he finds a situation such as that shown in Figure 5.6. Two heat reservoirs are indirectly connected by a closed gas-filled system. Thermal contact maintains each end of the gas system at the temperature of the adjacent reservoir. Thus, the gas at one end is hot, that at the other end is cold. The demon operates a partition which may be opened and closed. Because his faculties are "so sharpened" the demon can observe each molecule as it approaches the partition and determine its speed.

The demon sets out to pump heat from the cold reservoir to the hot reservoir. He does this as follows: The demon opens the partition whenever letting the molecule pass will raise the average molecular speed of the hot gas and lower the average molecular speed of the cold gas. There are two ways of doing this: (a) Let fast molecules excape from the cold gas, thereby cooling it and heating the hot gas. (b) Let slow molecules escape from the hot side, thereby increasing the average speed of those remaining, and thus also, raising the temperature. The demon in effect sorts and trades molecules. Equal numbers of molecules are passed in both directions so that the masses of the hot and

† J. C. Maxwell, "Theory of Heat," p. 338. Longmans, Green, New York (1894).

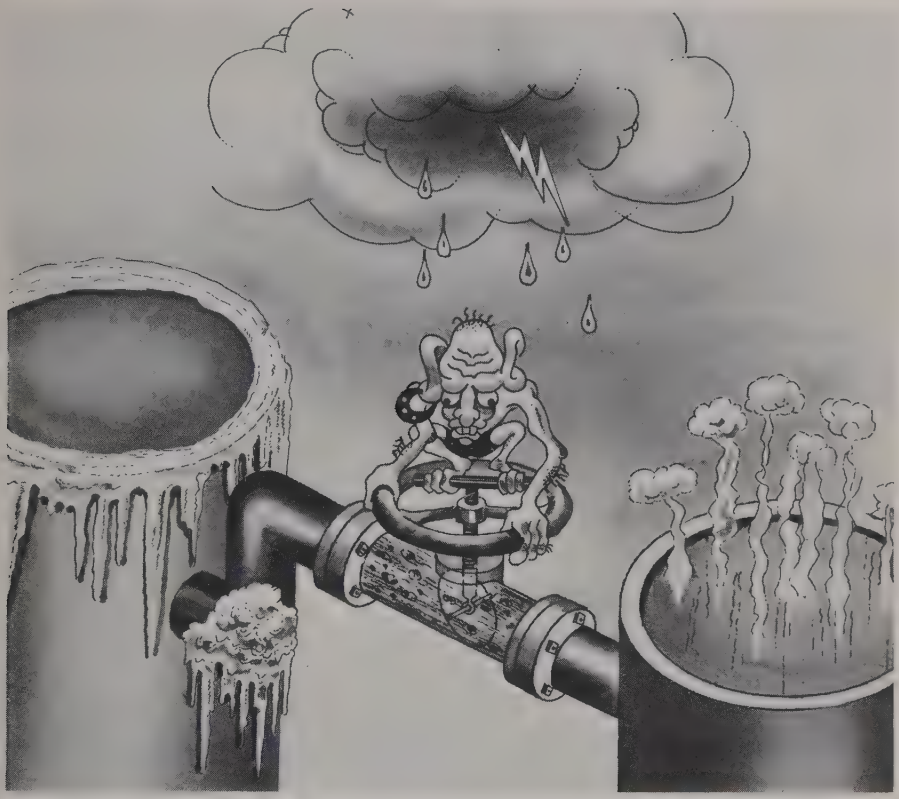


Figure 5.6. The demon controls the flow of molecules between the hot and cold reservoirs.

cold gases are not changed. Further, the volume containing the gases does not change so that no work is done by the gas. Now for the point! The action of the demon would lower the temperature of the cold gas were it not connected to a cold reservoir. The reservoir supplies whatever heat is required to maintain the temperature of the cold gas. Similarly, the hot gas rejects heat thereby maintaining its temperature at a constant value. The heat rejected by the hot gas might be used to power a heat engine.

The net effect of the demon's actions are to transfer heat from a cold reservoir to a hot reservoir. The gases on either side of the barrier are unchanged. Thus, we have a device—cyclic in nature—whose *only* effect is to transfer heat from a low temperature to a higher temperature. Such a demon could fry eggs with ice cubes, heat a home by extracting heat energy from the atmosphere, and run all the machinery, power plants, and so forth, on Earth by extracting

thermal energy from the oceans or the Earth itself. Unhappily, there is one minor problem—it will not work! Why not? Because we forgot the demon. We did not include the demon as part of the system. The demon cannot be regarded as part of the surroundings since he must communicate with the molecules in order to measure their speeds. The measurements made by the demon (or any instrumentalized version of the demon) require that he exchange energy with the molecules. Some of the energy being pumped uphill must be fed to the demon—he does not work free. Once the system is enlarged to include the demon one discovers that the heat pump has sprung a leak—a large leak. The demon is so hungry that he eats himself out of a job. His appetite completely wipes out the potential gain of his molecule-sorting action. To demonstrate this quantitatively is not easy. Basic to any detailed analysis of Maxwell's demon is the concept of entropy.† Entropy is a thermodynamic state function, a property of a system. A quantitative development of thermodynamics—not just the analysis of Maxwell's demon—necessitates the introduction of entropy. In the next chapter we study this essential thermodynamic state function from a macroscopic viewpoint. A microscopic view of entropy is forthcoming in Chapter 15.

REFERENCES

Carnot's paper is translated and interpreted in:

E. Mendoza (ed.), "Reflections on the Motive Power of Fire and Other Papers on the Second Law of Thermodynamics." Dover, New York (1960).

In particular, Mendoza notes in the Introduction that French physicists interpreted the caloric theory as meaning (a) heat is conserved and (b) heat and work are equivalent.

Attempts to circumvent the second law of thermodynamics by means of fluctuations which drive unilateral devices have been advanced from time to time. One of these devices—the ratchet and pawl—is dissected and demolished by:

R. P. Feynman, R. B. Leighton, and M. Sands, "The Feynman Lectures on Physics," Volume 1, Lecture 46. Addison-Wesley, Reading, Massachusetts (1963).

For details of the Stirling engine see:

M. W. Zemansky, "Heat and Thermodynamics," 5th ed., Chapter 7. McGraw-Hill, New York (1968).

† Analyses of Maxwell's demon from an entropy viewpoint are listed in the References at the end of this chapter.

A versatile demonstration model of the Stirling engine (operable as engine or refrigerator) is available from LaPine Scientific Co., Chicago, Illinois.

A vivid picture of what a skillful demon could accomplish is presented in:

G. Gamow, "Mr. Tompkins in Paperback," Chapter 9. Cambridge Univ. Press, London and New York (1971).

Maxwell's demon is discussed at some length by:

W. Herenberg, Maxwell's Demon, *Sci. Am.* **217**, 103 (Nov. 1967).

A historical survey of Maxwell's demon is presented in:

L. Brillouin, "Science and Information Theory," 2nd ed., Chapter 13. Academic Press, New York (1962).

Chapter 6

Entropy

In this chapter we introduce *entropy*, a thermodynamic state function. The introduction of entropy is vital to the quantitative development of thermodynamics. In particular, we can state the second law of thermodynamics in a most succinct fashion in terms of entropy. In Sections 6.1 and 6.2 we concern ourselves with "discovering" entropy and with seeing how it aids the mathematical development of thermodynamics. The entropy formulation of the second law is presented in Section 6.3. The chapter concludes with a discussion of the role of the second law in arguments which predict a "heat death" for the universe.

6.1 ENTROPY AND THE CARNOT CYCLE

In the last chapter, we discussed the Carnot cycle, a series of reversible processes which involve heat transfer. Recall that a reversible process is one which may be reversed by an infinitesimal change in the surroundings. A system and its surroundings may exchange heat reversibly if their temperatures differ infinitesimally. To establish this fact we first observe that heat flows into or out of a system depending upon whether its temperature is less than or greater than the temperature of the surroundings. The heat exchange is reversible only so long as an infinitesimal change in the temperature of the

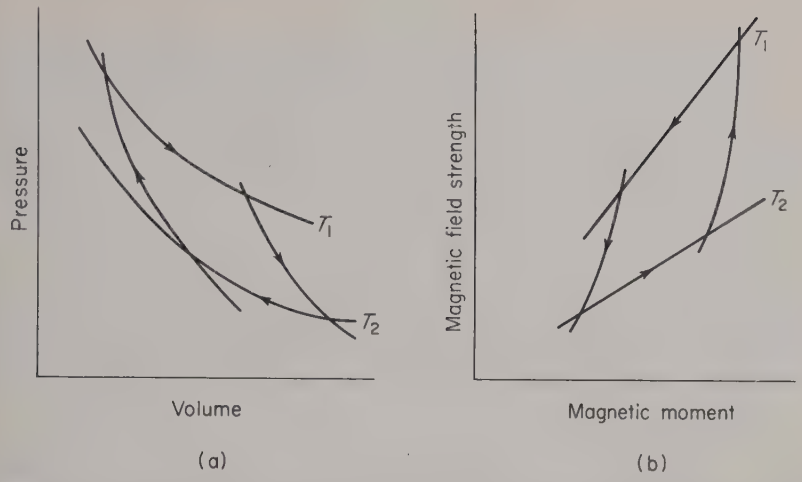


Figure 6.1. Carnot cycles for (a) an ideal gas and (b) for a paramagnetic solid. For the sequence of steps indicated by the arrows, both cycles result in the absorption of heat Q_1 by the system at a temperature T_1, and the rejection of heat Q_2 at a lower temperature T_2.

surroundings can reverse the flow of heat. However, this is possible only if the temperatures of the system and its surroundings differ infinitesimally. Figures 6.1a and b show the steps in Carnot cycles as executed by an ideal gas and a paramagnetic substance.

....

6.1.1 Refer to Figure 6.1. Both Carnot cycles absorb heat at the higher temperature T_1 and reject heat at the lower temperature T_2. Both convert heat into work. Explain why the P–V and H–M cycles are traced in opposite senses.

....

The quantities of heat absorbed and rejected along the isotherms of the Carnot cycle are related to the corresponding temperatures, namely,

$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2} \tag{6.1}$$

In Chapter 5 we temporarily repealed our sign convention on heat; both Q_1 and Q_2 are positive quantities in (6.1). We now reinstate our earlier convention that heat absorbed by the system is positive and that heat rejected by the system is negative. In (6.1), Q_2 denotes heat rejected so that we should now write

$$\frac{Q_1}{T_1} = \frac{-Q_2}{T_2} \qquad (Q_2 < 0)$$

or

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0 \tag{6.2}$$

There are four legs to the Carnot cycle: two isotherms and two adiabats. Because there is no heat exchange along the adiabats we can rewrite (6.2) as

$$\sum_{\text{all legs}} \frac{Q_i}{T_i} = 0 \tag{6.2a}$$

Furthermore, since the cycle must be traversed slowly to guarantee reversibility we can envisage the finite heat exchanges as being built up as sequences of infinitesimal exchanges. The summation over the cycle is then replaceable by an integral, namely,

$$\oint \frac{dQ}{T} = 0 \tag{6.3}$$

Equation (6.3), which has been deduced for the Carnot cycle, is in fact valid for any reversible process. A general proof is unnecessarily involved. For our purposes it suffices to consider Figure 6.2, which suggests how the result of (6.3) is established for a fluid system for which P, V, and T serve as the thermodynamic variables. The arbitrary cyclic process is mentally dissected into a succession of Carnot cycles. Each segment of the cycle path forms one of the isothermal legs of a Carnot cycle. Traversing all of the Carnot cycles once is equivalent to traversing the actual cycle path once. This is because each interior adiabat forms a leg of two adjacent Carnot paths. All such legs are

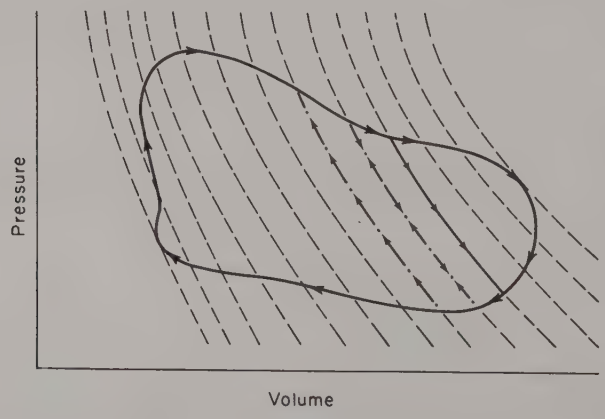


Figure 6.2. An arbitrary cyclic path may be chopped into a series of equivalent Carnot cycles to show that $\oint dQ/T = 0$ for a reversible process. The dashed lines represent the adiabats.

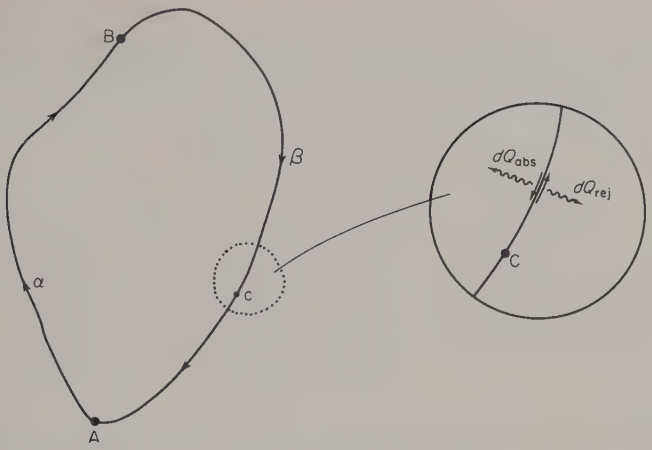


Figure 6.3. Symbolic representation of a reversible cyclic process, leading from A to B along path α and from B to A along path β. Heat absorbed (rejected) during each step would be rejected (absorbed) if the process were reversed.

traversed twice, in opposite directions, and in effect cancel out of the analysis. Since (6.3) applies to each Carnot cycle individually it is also valid for them collectively, and this in turn suggests that it holds for all reversible processes.

As a consequence of (6.3) we can infer the existence of a state function, called the entropy. To see how such an inference is made, consider a cyclic reversible process which carries a system through a pair of thermodynamic states labeled A and B. We can think of A and B as two points joined by a pair of paths (compare Figure 6.3). We can write the integral of (6.3) as a sum of two integrals. Thus

$$\oint \frac{dQ}{T} = \left(\int_A^B \frac{dQ}{T}\right)_\alpha + \left(\int_B^A \frac{dQ}{T}\right)_\beta$$

which may be reexpressed as

$$\left(\int_A^B \frac{dQ}{T}\right)_\alpha = \left(\int_B^A \frac{-dQ}{T}\right)_\beta \tag{6.4}$$

Now, because the path is reversible, we can write for the integral on the right side

$$\left(\int_B^A \frac{-dQ}{T}\right)_\beta = \left(\int_A^B \frac{dQ}{T}\right)_\beta \tag{6.5}$$

Physically, (6.5) follows because if heat dQ is absorbed along some segment as the system proceeds from A to B along β, an equal amount will be rejected along the same segment as the system moves from B to A. Mathematically

(6.5) may be thought of as a special case of (6.4) where α and β denote the same path. Comparing (6.4) and (6.5) shows that the integral of dQ/T between A and B is the same for both paths,

$$\left(\int_A^B \frac{dQ}{T}\right)_\alpha = \left(\int_A^B \frac{dQ}{T}\right)_\beta \tag{6.6}$$

The only restriction on α and β is that they be reversible paths connecting A and B. Equation (6.6) says that the value of the integral is independent of the reversible path from A to B. If one "walks" from A to B and adds up dQ/T along each step, then, no matter which reversible path one takes, the total is the same. We can therefore drop the subscripts which designate different paths. The value of the integral *does* depend on the states A and B and we can write

$$\int_A^B \frac{dQ}{T} = \int_A^B dS = S(\mathrm{B}) - S(\mathrm{A}) \tag{6.7}$$

where

$$dS = \frac{dQ}{T} \tag{6.8}$$

is the *entropy change* resulting from the reversible exchange of heat dQ with a source of temperature T. The entropy S is a thermodynamic state function—a property of the system—implicitly defined for all equilibrium states. From the relation $dS = dQ/T$ we see that the absorption of heat ($dQ > 0$) produces an entropy increase ($dS > 0$). It is important to remember that (6.8) holds only for a reversible process.

....

6.1.2 The heat capacity $C(T)$ for a system is given by

$$C(T) = AT^3$$

where A is a constant. If the entropy of the system is taken to be zero at $T = 0$, integrate (6.8) to show that the entropy at temperature T is

$$S(T) = \tfrac{1}{3}AT^3 = \tfrac{1}{3}C(T)$$

....

The units of entropy are those of energy divided by temperature; for example, calories per kelvin. The abbreviation eu stands for entropy unit, where one entropy unit is defined as one calorie per kelvin,

$$1 \text{ eu} \equiv 1 \text{ cal/K}$$

For reversible processes we can use (6.8) to express dQ as

$$dQ = T\,dS \tag{6.8a}$$

This is a noteworthy result. It expresses the heat absorbed solely in terms of the thermodynamic coordinates of the system. No reference is made to the surroundings. A similar situation arises in connection with the work dW performed by a system. Only for reversible processes is it possible to express dW in terms of

$$dW = P\,dV \tag{4.4}$$

or

$$dW = -H\,dM \tag{4.15}$$

Using (6.8a) and (4.4), the first law becomes

$$T\,dS = dU + P\,dV \tag{6.9}$$

a form suited only to a description of reversible processes.

Entropy is an extremely important physical property. It is probably also unfamiliar. Our immediate experience does not seem to contain the entropy concept. In fact, however, entropy is known to us all, but under other names. Entropy, it turns out, is the scientific alias for the familiar idea of *disorder*. The next section is devoted to exposing entropy as disorder under another name, and to presenting examples of entropy computations.

6.2 ENTROPY—EXAMPLES AND INSIGHTS

The concept of entropy will remain vague until we have an answer to the question, "What does entropy measure?" We shall emphasize the concept of entropy which has evolved from the microscopic domains of kinetic theory and statistical mechanics. Ludwig Boltzmann was the first to emphasize that the entropy of a system is a measure of its randomness or disorder. Disordering at the molecular level registers macroscopically as an entropy increase.

In this section we present examples which demonstrate how changes in entropy are computed. The examples selected illustrate the idea that entropy measures disorder.

Example 6.1 Entropy, Phase Changes, and Disorder

Consider a 1-gm ice cube at its melting point. If we add the latent heat of fusion, about 80 cal,[†] the ice cube melts. This change of phase takes place at

[†] The various latent heats vary with pressure. The numerical values quoted refer to a pressure of 1 atm.

constant temperature. The process is reversible so that we may integrate

$$dS = \frac{dQ}{T} \tag{6.8}$$

to obtain

$$\Delta S_{sl} = \frac{L_f}{T} \tag{6.10}$$

where

$$\Delta S_{sl} \equiv S_{\text{liquid}} - S_{\text{solid}} \tag{6.11}$$

is the entropy change, and L_f is the latent heat of fusion, the heat required to melt 1 gm of the solid. In the case at hand, the latent heat of fusion is approximately 80 cal and the melting point is roughly 273 K, whence,

$$\Delta S_{sl} \simeq \frac{80}{273} \frac{\text{cal}}{\text{K}} = 0.29 \text{ eu}$$

First we observe that ΔS_{sl} is positive—melting results in an entropy increase. Secondly we point out that the molecular structure of the final state (liquid) is highly randomized by comparison with the initial state (solid). In a crystalline solid the molecules form an ordered periodic structure. In a liquid only a short-range order is evident. By short-range order, we mean that any given molecule is only able to influence the behavior of nearby molecules. Melting results in an entropy increase and molecular disordering. Entropy measures disorder.

The computed value, $\Delta S_{sl} = 0.29$ eu, standing alone, is rather barren. To provide a basis for comparison we next compute entropy changes for two other processes involving the gram of water; (a) raising its temperature isobarically to the boiling point and (b) vaporizing it.

If heat is added to the liquid at constant pressure, its temperature rises until it reaches the boiling point. The heat added in changing the temperature by an amount dT is

$$dQ = C_p \, dT \tag{6.12}$$

where C_p is the specific heat at constant pressure. The corresponding entropy change is

$$dS = \frac{C_p \, dT}{T} \tag{6.13}$$

For many liquids, and water in particular, C_p is nearly constant between the melting and boiling points. If C_p is treated as constant, the total entropy

change between the melting and boiling points is

$$\Delta S_{\text{mb}} = C_p \int_{T_{\text{m}}}^{T_{\text{b}}} \frac{dT}{T} = C_p \ln\left(\frac{T_{\text{b}}}{T_{\text{m}}}\right) \tag{6.14}$$

For 1 gm of water, $C_p \simeq 1$ cal/K and $\ln(T_{\text{b}}/T_{\text{m}}) \simeq 0.31$. These data give

$$\Delta S_{\text{mb}} \simeq 0.31 \text{ eu}$$

This is only slightly larger than the entropy change of melting. Raising the temperature of the water increases the molecular disorder. This additional randomization is due largely to the fact that the average molecular speed increases with temperature. As the liquid is heated, molecular agitation grows. The chance that a given molecule influences the behavior of a nearby molecule decreases as their relative speed increases—they pass too swiftly to communicate effectively. Thus, as the temperature increases, molecular correlations diminish and disorder rises.

Having raised the water to its boiling point, it can be vaporized by adding still more heat. The entropy change in going from the liquid to the vapor phase is readily found by the same arguments leading to (6.11). The result is

$$S_{\text{vapor}} - S_{\text{liquid}} \equiv \Delta S_{lv} = \frac{L_{\text{v}}}{T} \tag{6.15}$$

where L_{v} is the latent heat of vaporization and T is the boiling point. For 1 gm of water at 1 atm, $L_{\text{v}} = 539$ cal, $T = 373$ K, and $\Delta S_{lv} \simeq 1.44$ eu. The entropy change for the liquid–vapor transition is roughly five times that of the two previous processes. What disruption at the molecular level can be identified as the cause of this entropy increase? What sort of disordering has occurred? A partial answer is to be found by noting that at the normal boiling point, 1 gm of liquid water occupies a volume of about 1 cm^3. At the same pressure and temperature the vapor occupies a volume of approximately 1700 cm^3. The same collection of molecules which started as a liquid is enormously spread out as a vapor. The molecules are greatly disorganized by virtue of having a larger volume in which to roam. It is somewhat like the disorder which would result if the door to a chicken coop were left open, permitting the chickens freedom to move about in a barnyard area 1700 times as large as that of the coop.

The entropy change in the liquid–vapor transition is not wholly accounted for by the increase in volume. There is also a contribution stemming from the breaking of molecular bonds. A molecule which escapes from the liquid to take its place in the vapor leaves behind a less ordered structure. There is a "hole" in the liquid. Such defects represent increased disorder and contribute to the total entropy change. ✧

•••••• ••••••

6.2.1 At a pressure of 1 atm, liquid helium boils at 4.2 K. The latent heat of vaporization is 4.9 cal/gm. Determine the entropy change (per gram) resulting from vaporization. Express your result in entropy units per gram.

6.2.2 The specific heat of copper (at constant pressure) is 0.092 cal/gm-°C.

(a) If the temperature of 20 gm of copper is raised from 0 °C to 20 °C, determine the entropy change of copper. Express your results in entropy units.

(b) How much heat is absorbed by the copper in the process?

(c) If all of the heat computed in part (b) were absorbed by the copper at the average temperature (10 °C), what would be the entropy change?

6.2.3 The latent heat of fusion of lead (Pb) is 1.12 kcal/mole. The normal melting point is 327.5 °C. Determine the entropy change of 100 gm of lead when it melts at a pressure of 1 atm.

•••••• ••••••

We have stated that S is a thermodynamic state function. The definition $dS = dQ/T$ does not tell us how S depends on the thermodynamic variables T, P, V, and so on. The functional form of S can be determined using the first law once we know (a) the equation of state and (b) the functional form of the internal energy. Example 6.2 demonstrates how the first law may be used to establish the form of S. Problem 6.2.4 gives the student an opportunity to try his hand at determining the entropy of a photon gas.

Example 6.2 Determination of the Entropy Using the First Law

The internal energy of a hypothetical substance is given by

$$U = aVT^2, \qquad a = \text{constant}$$

and its equation of state is

$$PV = U$$

or

$$P = aT^2$$

We wish to determine the entropy of the system as a function of T and V. Starting with the relation

$$dS = \frac{dQ}{T}$$

we use the given information and the first law to express dQ in terms of dT and dV. This expression for dQ leads to a form for dQ/T which then must be recognized as the differential of some explicit function of T and V. From the

equation of state we find

$$P\,dV = aT^2\,dV$$

Forming the differential of the expression for U gives

$$dU = aT^2\,dV + 2aVT\,dT$$

Inserting these forms of dU and $P\,dV$ into the first law gives dQ, which in turn leads to

$$dS = \frac{dQ}{T} = 2aT\,dV + 2aV\,dT$$

The right-hand side must be recognized as the differential of $2aTV$

$$dS = d(2aTV)$$

The functional form of S is therefore

$$S(T,V) = 2aTV + S_0$$

where S_0 is a constant of integration.

In the next section we use this technique to determine the entropy of an ideal gas. ✧

....

6.2.4 The internal energy of a photon gas is given by

$$U = bVT^4$$

where b is a constant. The equation of state is

$$PV = \tfrac{1}{3}U$$

Derive an expression for the entropy of the photon gas. Choose the constant of integration to be such that the entropy is zero at $T = 0$.

....

6.3 THE IDEAL GAS, FREE EXPANSION, AND GIBBS' PARADOX

An expression for the entropy of an ideal gas is readily derived and proves useful in many applications. Consider 1 mole of an ideal gas. Using the first law we can write

$$dS = \frac{dQ}{T} = \frac{dU + P\,dV}{T}$$

For an ideal gas,

$$dU = C_v\,dT$$

and

$$PV = RT$$

The heat capacity C_v is a constant. The expression for dS converts to

$$dS = C_v \frac{dT}{T} + R\frac{dV}{V}$$

This relation may be integrated to give

$$S = C_v \ln T + R \ln V + S_0 \tag{6.16}$$

where S_0 is an integration constant. The value of S_0 is of no interest so long as we concern ourselves only with entropy changes. For example, in a process which carries 1 mole of an ideal gas from (T_1, V_1) to (T_2, V_2), the entropy change is

$$S_2 - S_1 = C_v \ln\left(\frac{T_2}{T_1}\right) + R \ln\left(\frac{V_2}{V_1}\right) \tag{6.17}$$

The unknown constant S_0 has dropped out.

6.3.1 Equation (6.16) gives the entropy for 1 mole of an ideal gas. Carry through the corresponding analysis to show that the entropy of n moles is simply n times that of 1 mole, that is, verify that entropy is an extensive quantity.

There are two rather instructive applications of (6.17). These are the free expansion of a gas and Gibbs' paradox. Consider the situation schematicized in Figure 6.4. A vessel with rigid, insulated walls is divided by a partition.

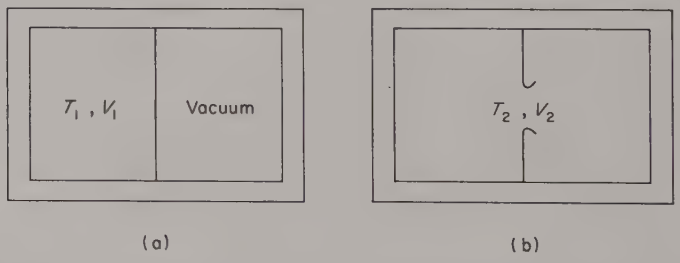


Figure 6.4. Initially, (a) one portion of the rigid, insulated container is a vacuum while an ideal gas occupies the other section. After the partition ruptures, (b) the gas expands to fill the container. The insulation and rigid walls guarantee that the system is isolated.

Initially, one side contains an ideal gas while the other side is a vacuum. Consider what happens when the partition ruptures, allowing the gas to expand and fill the entire container. The insulation prevents heat exchange. Thus, $dQ = 0$ for each stage of the expansion. The rigid walls guarantee that no work is done on the system, that is, there can be no change in the volume of the system and thus no $P\,dV$ work by the surroundings. The process is called a free expansion because of the absence of any external work as the gas rushes (or oozes) into the evacuated region. We therefore have the results that $dQ = 0$ and $dW = 0$ at each stage of the expansion. It follows from the first law that $dU = 0$; the internal energy of the gas does not change. However, for an ideal gas $U = C_v T$, the internal energy depends on T alone. If U remains constant, so must the temperature; thus $T_2 = T_1$. The final volume of the gas V_2 is clearly greater than the initial volume V_1, so that $\ln(V_2/V_1)$ is a positive number. From (6.17) we then have the result that

$$S_2 - S_1 = R \ln\left(\frac{V_2}{V_1}\right) > 0$$

that is, the final entropy exceeds the initial entropy. The gas has expanded but its temperature has not changed; the gas spreads out over a larger volume, resulting in further disorder. The predicted entropy increase concurs with the notion that entropy measures disorder. Very straightforward—or is it? Let us retreat momentarily to the point where we prescribed that the container was insulated. We noted that the insulation guaranteed that $dQ = 0$ at each stage of the expansion. From

$$dS = \frac{dQ}{T} \tag{6.8}$$

it might seem that we should conclude that $dS = 0$ at each stage, and therefore that the entropy of the gas remains constant. Not so! The entropy does change. The expression $dS = dQ/T$ does not apply to processes such as the free expansion. The free expansion is an irreversible process. It cannot be reversed by any infinitesimal change in the surroundings. As we warned earlier the relation $dS = dQ/T$ holds only for reversible transformations.

Students are apt to encounter two types of paradoxes. On the one hand, there are the pseudo-paradoxes. These are pedagogically designed and motivated. By resolving them the student gains insight and understanding—he's educated! The various perpetual motion machines suggested earlier are examples of such paradoxes. They are hoaxes; understanding precisely why they do not work is instructive.

The second type of paradox can lay some claim to legitimacy. At some point in history it was a genuine paradox—a facet of nature which could not be explained. Historically, the resolution of such paradoxes have brought about

modifications of existing physical laws. Gibbs' paradox is a case in point. Gibbs' paradox is no longer paradoxical. It can be understood—explained—via quantum statistical mechanics. Consider again the insulated container of Figure 6.4. In the initial equilibrium state, one section contains ideal gas A, the other section contains ideal gas B. For simplicity we suppose that the gases have equal temperatures and that they occupy equal volumes. The initial entropy of the system is the sum of the entropies of A and B.

$$S_i = C_{vA} \ln T + C_{vB} \ln T + 2R \ln V + S_{0A} + S_{0B}$$

The partition separating the two gases is now removed, permitting the two gases to mix. In the final equilibrium state the now-mixed gases are at temperature T (the temperature does not change) and occupy a volume $2V$, the total container volume. The final entropy differs from the initial entropy only in that both A and B occupy a volume $2V$ rather than V. Thus

$$S_f = C_{vA} \ln T + C_{vB} \ln T + 2R \ln 2V + S_{0A} + S_{0B}$$

The entropy change

$$S_f - S_i = 2R \ln 2V - 2R \ln V = 2R \ln 2 \tag{6.18}$$

is positive. The entropy increases because the gases have mixed; mixing results in greater disorder. In fact, the entropy increase computed above is referred to as the entropy of mixing.

So far, there is no paradox! The result is in accord with the idea that entropy measures disorder. However, Josiah Willard Gibbs made the following observation, from which springs the paradox: Suppose that gases A and B are *identical*. The initial and final states of the system are then identical. There is no operational way of detecting any mixing, for none has occurred. If this point is not clear, mentally replace the atoms with marbles. Suppose first that the marbles are distinguishable by their color. If the A marbles are white and the B marbles are black, any mixing is readily detected. However, if the marbles are all black, no mixing can be detected even though marbles change their positions. Regardless of their physical nature, so long as the A and B objects are identical no experimental test can distinguish between the initial and final states.

Now for the paradox! When A and B are identical there is no mixing; the initial and final states are identical. Hence there is no entropy change. The paradox, first pointed out by Gibbs, is that the entropy change predicted by (6.18) is the same ($2R \ln 2$) even when A and B are identical gases. The very straightforward calculation performed earlier predicts an increase in entropy even when the gases are identical—when in fact there is no entropy change. Something is wrong! There were no mathematical errors. We have a genuine paradox. A correct result is obtained so long as A and B are different. It is the

physics which has gone wrong. The expression for the entropy of an ideal gas must be modified in some way to account for the discrepancy. The modification should be such that the entropy of mixing becomes zero when the gases are identical.

The correct expression for the entropy was deduced by Gibbs. However, his was an *ad hoc* recipe. He saw how to obtain a satisfactory answer but could not justify his procedure. Little wonder! The fundamental basis for the corrected entropy formula lies within modern quantum theory, a discipline that was to develop 30 years after Gibbs' insight led him to the correct expression. Problem 6.3.4 presents the corrected entropy formula and an invitation to verify its superiority.

....

6.3.2 Explain why no temperature change occurs during the expansion and mixing described in connection with Gibbs' paradox.

6.3.3 Two vessels, both containing 1 mole of an ideal gas, are brought into thermal contact. The initial temperatures of the gases are 300 and 450 K. The gases are identical, and occupy equal volumes which remain constant.

(a) What is their final common temperature?
(b) Compute the entropy change.
(c) Optional: Consider the more general case where the initial temperatures are T_1 and T_2 and prove that the entropy always increases provided $T_1 \neq T_2$. [*Hint:* The arithmetic mean of two unequal numbers is greater than the geometric mean $\frac{1}{2}(x+y) > \sqrt{xy}$.]

6.3.4 Gibbs observed that his paradox was resolved if the entropy for n moles of an ideal gas is written as

$$S_{\text{cor}} = nC_v \ln T + nR \ln\left(\frac{V}{n}\right) + nS_0$$

rather than as

$$S = nC_v \ln T + nR \ln V + nS_0$$

The two expressions differ in that S_{cor} (corrected) has V/n, the volume per mole, in place of V. When A and B are identical the volume per mole of gas remains unchanged in the mixing experiment described earlier. Verify that (6.18) again results when A and B are different (using S_{cor}), but that there is no entropy change when A and B are identical. Take $n_A = n_B = 1$ mole for convenience.

....

6.4 *T–S* DIAGRAMS

We have seen how a *P–V* diagram may be used to interpret the work done by a fluid system. A temperature–entropy (*T–S*) diagram serves a similar purpose for the heat absorbed or rejected in a process. Suppose a system undergoes a reversible transformation during which the temperature and entropy vary as shown in Figure 6.5. From the figure we see that the heat

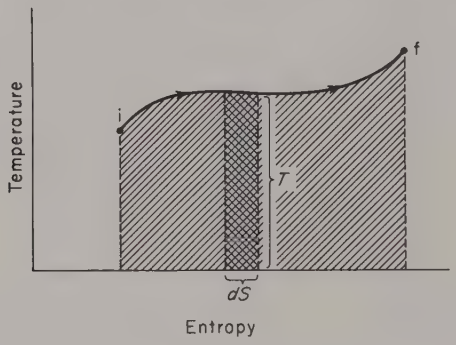


Figure 6.5. A reversible transformation takes the system from i to f. The solid line indicates the relation between T and S during the process. The cross-hatched area is $T\,dS$. The total hatched area beneath the solid line corresponds to the heat absorbed by the system during the process.

absorbed by the system can be given a geometrical interpretation. For an entropy change dS, the heat absorbed is

$$dQ = T\,dS \tag{6.8a}$$

As the figure shows, $T\,dS$ corresponds to the cross-hatched area of the strip of height T and width dS. The hatched area in Figure 6.5 corresponds to the total heat absorbed during the process

$$Q = \int_{\mathrm{i}}^{\mathrm{f}} T\,dS$$

Figure 6.6a shows the *T–S* diagram for a Carnot cycle. The two adiabatic legs of a Carnot cycle are also isentropic because the process is reversible.† Figure 6.6b shows the corresponding steps on a *P–V* diagram.‡ During the

† Recall that $dS = dQ/T$ applies to reversible processes. If $dQ = 0$ (adiabatic), then $dS = 0$ as well (isentropic).

‡ The *P–V* diagram shows a Carnot cycle for an ideal gas. The *T–S* plot is universal in that its shape is the same regardless of the working substance undergoing the Carnot cycle.

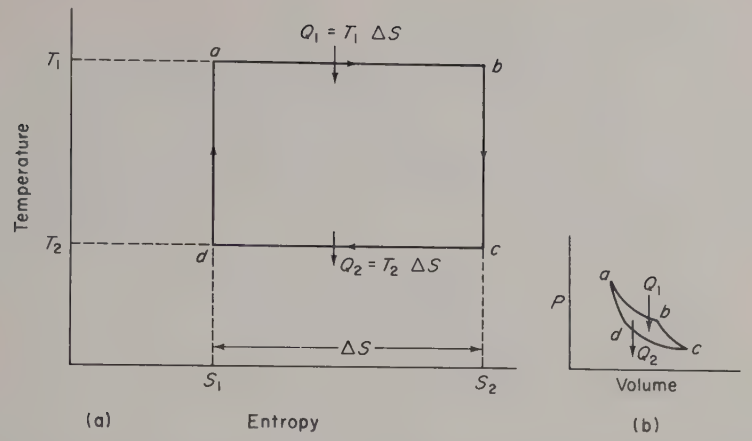


Figure 6.6. (a) The *T–S* diagram for the Carnot cycle. (b) The corresponding *P–V* diagram for an ideal gas.

isothermal transition from a to b an amount of heat

$$Q_1 = T_1 \Delta S$$

is absorbed by the system. Geometrically, $T_1 \Delta S$ is the area of the rectangle of height T_1 and width ΔS. Along the isotherm from c to d the system rejects heat of magnitude

$$Q_2 = T_2 \Delta S$$

which also corresponds to a rectangular area. The net amount of heat absorbed by the system is $Q_1 - Q_2$. From Figure 6.6 we see that this corresponds to the rectangular area generated by the cyclic path on the *T–S* diagram. Observe also that the cycle path is traced clockwise and that the system absorbs a net amount of heat. If the Carnot cycle were reversed, the enclosed area would then correspond to the net heat rejected to the surroundings.

It follows from the first law that $W = Q_1 - Q_2$ is the work done by the system during the cycle, that is, the enclosed area represents the work done by the system. We have used the Carnot cycle because it is a familiar example. The foregoing statements remain valid for all reversible processes.

In the case of the Carnot cycle, the *T–S* diagram leads to an instant derivation of its thermodynamic efficiency. From its definition,

$$\eta = \frac{\text{work performed}}{\text{heat absorbed}} = \frac{W}{Q_1}$$

From the *T–S* diagram,

$$W = Q_1 - Q_2 = T_1 \Delta S - T_2 \Delta S$$

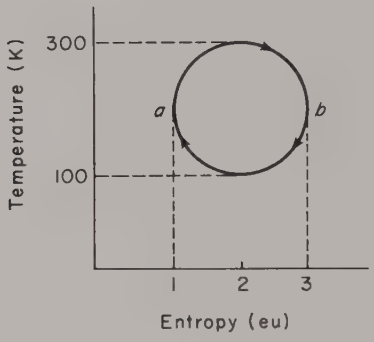
Inserting this gives the desired result

$$\eta = \frac{(T_1 - T_2)\Delta S}{T_1 \Delta S} = 1 - \frac{T_2}{T_1}$$

....

6.4.1 Show that the thermodynamic efficiency of the circular cycle shown in the accompanying figure is

$$\eta = \frac{100\pi}{400 + 50\pi}$$



....

Before moving on to an entropy formulation of the second law of thermodynamics, two remarks are appropriate. First, our insistence that entropy measures disorder has, at best, been made plausible. Little more could be expected, since we have not attempted to define disorder. A quantitative approach to the disorder–entropy relation is possible only at the microscopic level. In so far as the formal manipulations and applications of thermodynamics are concerned, we may regard entropy as just another useful physical property. Secondly, we have quietly, and purposely, ignored the question of how entropy is determined experimentally. This is related to another point swept under the rug—the entropy constant which appeared in (6.16). In principle the measurement of entropy change is quite straightforward. By adding or extracting heat dQ reversibly, the entropy change follows from

$$dS = \frac{dQ}{T}$$

Until a value of entropy is assigned to some standard state we cannot tabulate anything beyond entropy changes. Why not just pick some convenient state and assign a value of (say) zero to the entropy of a compound in that state?

The answer is that nature will not let us pick just any state. Experiment shows that the entropy of all systems approach a common value as the temperature tends toward absolute zero. This is the gist of the third law of thermodynamics. (We shall state it more precisely later.) The practical problem of tabulating entropies therefore involves the third law, and the behavior of substances at temperatures approaching absolute zero. We will meet these problems head-on in the next chapter.

6.5 ENTROPY FORMULATION OF THE SECOND LAW OF THERMODYNAMICS

In Section 6.3 we studied two processes involving isolated systems; the free expansion of a gas and the mixing of two gases. Both resulted in an entropy increase, and both are irreversible processes. These two processes exemplify an alternate statement of the second law of thermodynamics, advanced by Rudolph Clausius:

The entropy of an isolated system never decreases.

This statement is generally referred to as the *entropy principle*. Stated mathematically this version of the second law reads,

$$\Delta S_{\text{I.S.}} \geqslant 0 \tag{6.19}$$

where ΔS denotes the change in entropy brought about by any thermodynamic processes which connect equilibrium states and I.S. is isolated system. If an isolated system undergoes a reversible process, $dS = dQ/T = 0/T = 0$; there is no change in its entropy. This statement is contained in the equality portion of (6.19). If an isolated system is not in equilibrium, irreversible processes operate to drive it toward equilibrium. Irreversible processes generate entropy. They mix, disorder, randomize, and so on. The inequality in (6.19) refers to nonequilibrium situations in which irreversible processes generate entropy. From these remarks it follows that equilibrium corresponds to a state of maximum entropy—the entropy increases until the system attains equilibrium, after which the entropy remains constant.

Care must be exercised in applying the entropy principle. An increase in entropy is claimed only for an isolated system. A healthy 10-year-old boy converts mashed potatoes into muscle; disorder into order, but the 10-year-old is not an isolated system. If any living system is isolated, its entropy increases. Disorder mounts; the system withers and dies.

On the grounds that thermodynamics can be fun if it is not made laborious we shall not present a mathematical proof of the entropy principle. The

interested student should consult the references. Instead we turn our attention to a question which has intrigued scientists for over a century. Does the second law of thermodynamics apply to the largest isolated system—the universe?

6.6 ENTROPY AND UNIVERSAL ACCIDENTS

The entropy principle, as formulated by Rudolph Clausius, states that the entropy of an isolated system never decreases. Entropy increases until it reaches a maximum value. The irreversible processes responsible for the increase of entropy are the same mechanisms which drive the system toward equilibrium. Thus, thermodynamic equilibrium corresponds to a state of maximum entropy.

The entropy principle is sometimes advanced in arguments for a "heat death" of the universe. The irreversible processes of nature result in an increase of entropy in the universe. The ultimate state of maximum entropy is one of thermodynamic equilibrium—a universe of uniform temperature.

However, the myriad of irreversible processes which perform so ably in establishing equilibrium in subcosmic chunks of the universe seem to lose their punch when applied to that part of the universe visible to us. The astronomers, astrophysicists, cosmologists, and anyone with one good eye can give evidence that the portion of the universe in our immediate neighborhood is not in equilibrium, nor are there any signs that it is tending toward equilibrium. This situation is a sharp contrast to those we normally observe, in which an isolated system rather quickly reaches equilibrium. The universe is certainly isolated—why does it not tend toward equilibrium?

One suggested resolution of this paradox is known as the fluctuation hypothesis. The fluctuation hypothesis was suggested by Ludwig Boltzmann, one of the fathers of statistical mechanics. All systems in thermodynamic equilibrium exhibit fluctuations. A reasonably strong analogy exists between fluctuations and auto accidents. Imagine a busy highway, with the traffic moving in a regular fashion. Occasionally two cars collide. Accidents involving three or more cars are less likely. Accidents involving many cars are very unusual—that is why newspapers report such incidents.

A fluctuation in a thermodynamic system is the result of an accident on the atomic level involving many atoms. In the atmosphere a density fluctuation occurs when many more (or many less) than the average number of molecules happen to populate some particular volume. A temperature fluctuation arises when a group of very fast (or very slow) molecules occupy some particular volume. Such fluctuations occur by chance—they are accidents of nature. For this reason, fluctuations within small volumes, which involve relatively few

particles, are more likely to occur than fluctuations over larger volumes, which necessarily involve many more particles. Big fluctuations are like big accidents—both are very unlikely.

Boltzmann's fluctuation hypothesis suggests that the universe is in equilibrium but that the portion which we observe is part of a gigantic fluctuation—the granddaddy of all accidents. At first sight, the argument for such a hypothesis seems to be a strong one. Some sort of fluctuation is required to ensure the existence of observers (you and me!), that is, biological development requires special conditions—conditions of a distinctly nonequilibrium nature. Thus the very fact that such biological development has occurred—that I write and you read—seems to be strong evidence for the fluctuation hypothesis. However, the argument is unsound. It is enormously more likely that such a fluctuation would occur over a small volume, say the size of our solar system, and leave the rest of the immediate universe in equilibrium. To pursue the traffic analogy, being involved in an accident is not unusual, but we would generally be able to see beyond the wreckage and discern the equilibrium flow of traffic. The chance that the fluctuation hypothesis is true is less than the likelihood of an accident involving every car on the road today. Such states of true chaos seem most unlikely.

Boltzmann's fluctuation hypothesis is untenable. How then can we explain the fact that we live in a restless universe? One possible answer is suggested by the theory of gravitation. This theory, better known as general relativity, deals with systems in variable gravitational fields. On a cosmic scale we can no longer assume a spatially uniform gravitational field. Further, we must allow for the time variation of gravitational effects. Our day to day observations of systems which tend toward equilibrium are restricted to situations in which the surroundings do not vary with time, or vary at a rate which is slow by comparison with the speed at which the system approaches equilibrium. The general theory of relativity teaches us that the geometry of space is determined by the masses within the space. As the distribution of mass in our universe varies so do the geometrical properties of space. The situation is roughly analogous to a gas in a container whose surface wiggles, twists, and dilates. Such a system would not approach equilibrium because the external conditions are changing. The variable nature of the gravitational forces within our universe makes it behave like a container whose volume wiggles, twists, and dilates. In a sense, our universe is not an isolated system.

You may ask, "What sort of thermodynamic history is predicted for our universe by the general theory of relativity?" The answer depends on what sort of *model* you pick for our universe. It also depends on the answer to a question astronomers are now trying to answer, namely, what is the average mass density of the universe? The present (estimated) value for the average mass density is 7×10^{-31} gm/cm^3. The observed expansion of the universe

will continue indefinitely unless the density is at least 2×10^{-29} gm/cm^3. Life in a forever expanding universe seems less attractive than in one which is "closed." For this and other (less psychologically motivated) reasons, astronomers are still looking for additional matter in the vast expanses of the universe.

REFERENCES

An interesting and enlightening extension of Gibbs' paradox is presented in:

M. J. Klein, Note on a Problem Concerning the Gibbs Paradox, *Am. J. Phys.* **26**, 80 (1958).

A careful and extensive development of the entropy concept, from both the classical and quantum viewpoints, is contained in:

J. S. Dugdale, "Entropy and Low Temperature Physics." Hutchinson, London (1966).

Chapter 13 includes comments on Maxwell's demon.

A rigorous proof of the entropy principle is given by:

E. Fermi, "Thermodynamics," Chapter 4. Dover, New York (1956).

Cosmological consequences of the "missing mass" are touched on in:

W. H. McCrea, Cosmology Today, *Am. Sci.* **58**, 521 (1970).

and

J. A. Wheeler, Our Universe: The Known and Unknown, *Am. Sci.* **56**, 1 (1968).

The origin of the irreversibility in nature is still a matter of debate. For a critical review, see:

B. Gal-Or, The Crisis about the Origin of Irreversibility and Time Aniosotropy, *Science* **176**, 11 (1972).

Chapter 7

The Third Law of Thermodynamics

7.1 ENTROPY, ABSOLUTE ZERO, AND THE THIRD LAW

While the remarks of this chapter aim at the third law, the reader should be reminded of the first two. The first law may be phrased as a warning[†]: "You cannot win. You can only break even." The first law is a statement of energy conservation. It recognizes the impossibility of winning with perpetual motion machines of the first kind—cyclic devices which would export more energy than they import. The second law is even more pessimistic about one's chances of winning. The second law recognizes the impossibility of constructing a cyclic device which converts heat completely into work. In terms of the thermal efficiency

$$\eta = \frac{\text{work output}}{\text{heat input}}$$

the second law may be stated

$$\eta < 1$$

For example, the efficiency of the Carnot cycle is

$$\eta_c = 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}}$$

[†] Compare, *Am. Sci.* **52**, 40A (1964).

It appears that if $T_{\text{cold}} = 0\,\text{K}$, then $\eta_c = 1$, and we would have a device which converts heat completely into work. This has led to the following one line statement of the second law: "You can break even only at absolute zero." However, is it possible to reach absolute zero? Experiment decides whether or not one can reach the state $T = 0\,\text{K}$. Experiment reveals that all schemes for lowering temperature suffer an attack of diminishing returns upon approaching absolute zero. They become less and less effective as the temperature decreases. The third law, which comes as a disappointment to those who aspire to break even may be stated as "You cannot reach absolute zero." The immediate conclusion to be drawn from the first, second, and third laws is that you can neither win nor break even. This makes thermodynamics a very expensive pastime.

There is an alternate way of stating the third law. It is phrased in terms of the entropy. The entropy interpretation of the third law reinforces the viewpoint stressed in Chapter 6—that entropy measures disorder.

The entropy change dS resulting from the reversible transfer of heat dQ at temperature T is

$$dS = \frac{dQ}{T}$$

This is an operational definition, not of entropy, but of the change in entropy. This did not strike us as unusual when it was first advanced because we have encountered quite similar situations before. For example, in mechanics only the change of potential energy is operationally defined. We are free to choose the position of zero potential energy at our convenience. In thermodynamics there is likewise a choice as to the state of zero internal energy. The first law defines only the change of internal energy.

Despite the fact that only dS is operationally defined we do not have complete freedom in the choice of the state of zero entropy. Nature rules otherwise. The third law of thermodynamics expresses this further restriction on the entropy. It was Walther Nernst who first saw the necessity of inventing a new law of thermodynamics. In 1907 Nernst advanced his heat theorem. Now enlarged, revised, and bolstered by additional experimental evidence, Nernst's Heat Theorem stands as the third law of thermodynamics. In its most unqualified form the third law is this:

> The entropy of a system approaches zero as its temperature tends toward absolute zero.

This statement must be qualified to make it precise and universally valid—to put it in the category of physical law. However, after all the qualifying "weasel" words are inserted, the essential content of the third law remains

$$S \to 0 \quad \text{as} \quad T \to 0$$

In the next section we cite evidence for the third law by showing how experiment confirms its predictions. In Section 7.3 we note certain exceptions, interpret them, and modify the original statement accordingly. In Section 7.4 we illustrate how the third law renders absolute zero unattainable. To fully appreciate the third law we must continue to cultivate the notion that entropy measures disorder. The statement that entropy tends to zero with temperature implies that absolute zero is the ultimate if not fully attainable state of order.

7.2 THE THIRD LAW IN ACTION

Let us accept the unrestricted statement of the third law, deduce some of its consequences, and then check to see if these are borne out by experiment. If we utilize one of the Maxwell relations,†

$$\left(\frac{\partial V}{\partial T}\right)_P = -\left(\frac{\partial S}{\partial P}\right)_T \tag{8.18}$$

the coefficient of volume expansion, $\beta = (1/V)(\partial V/\partial T)_P$, can be expressed as

$$\beta = -\frac{1}{V}\left(\frac{\partial S}{\partial P}\right)_T$$

Using the third law we can show that $(\partial S/\partial P)_T$ falls toward zero with the temperature. Therefore, the third law predicts that β tends toward zero as the temperature approaches absolute zero. To show that $(\partial S/\partial P)_T$ tends to zero we note that it corresponds to the experimentally observed quantity

$$\frac{\Delta S}{\Delta P} = \frac{S(T, P+\Delta P) - S(T, P)}{\Delta P}$$

where $\Delta S/\Delta P$ is a function of T and P. If the entropy falls to zero as T approaches zero, it follows that the entropy change ΔS also must tend to zero, regardless of the value of ΔP. Thus the ratio $\Delta S/\Delta P$ goes to zero

$$\frac{\Delta S}{\Delta P} = \left(\frac{\partial S}{\partial P}\right)_T \to 0 \qquad \text{as} \qquad T \to 0$$

and it follows that β tends to zero with the temperature. Figure 7.1 shows a plot of α ($\alpha = \frac{1}{3}\beta$) versus temperature for gold. The figure reveals that α approaches zero as the temperature tends toward absolute zero, in accord with the third law.

---

† Maxwell's relations, including (8.18), are derived in Chapter 8.

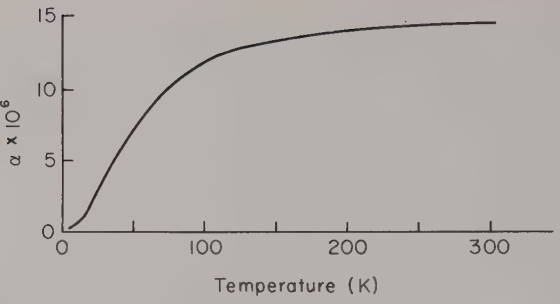


Figure 7.1. Coefficient of linear expansion α for gold. (After H. M. Rosenberg, "Low Temperature Solid State Physics." Oxford Univ. Press, London and New York (1963).

A second consequence of the third law concerns the difference of the heat capacities $C_p - C_v$. We noted earlier that C_p and C_v differ because materials tend to expand or contract when heated at constant pressure. However, as β goes to zero the volume ceases to change with the temperature. We therefore expect C_p and C_v to become equal as T approaches zero. In fact, examination of (4.26) shows that the difference of C_p and C_v is proportional to

$$\left(\frac{\partial V}{\partial T}\right)_P = \beta V$$

a quantity which vanishes as $T \to 0$. Figure 7.2 shows the low-temperature confluence of C_p and C_v for solid neon.

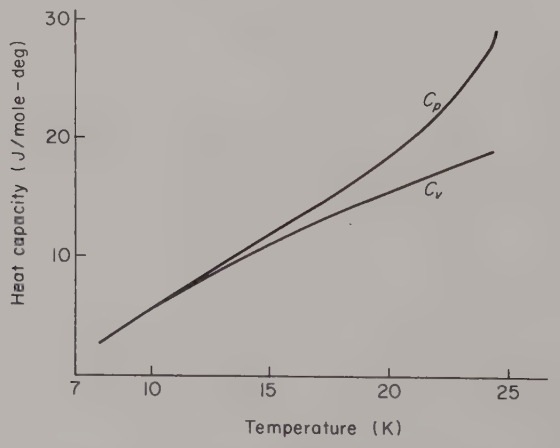


Figure 7.2. Heat capacities of solid neon. [After C. H. Fagerstroem and A. C. Hollis Hallet, "Low Temperature Physics," Part B. Plenum, New York (1965).]

A third experimental confirmation of the third law is provided by low-temperature heat capacity measurements. The third law not only predicts that C_p and C_v become equal, it also predicts that each vanishes as T falls toward absolute zero. To show this we write

$$dQ = C\,dT$$

where C is the appropriate heat capacity. The corresponding entropy change is

$$dS = \frac{C\,dT}{T} \tag{7.1}$$

If we integrate this equation from absolute zero to some temperature T, we obtain

$$S(T) - S(0) = \int_0^T \frac{C\,dT}{T} \tag{7.2}$$

At this point we seize an opportunity to qualify our original statement of the third law. A more precise statement is that the entropies of all systems approach the *same* value as the temperature tends toward zero. The choice of this common value is left open, and the scientific community has settled on zero. With the choice $S(0) = 0$, the integral in (7.2) fixes the entropy at temperature T as

$$S(T) = \int_0^T \frac{C\,dT}{T} \tag{7.3}$$

The behavior of C as T approaches zero is severely restricted by the condition that

$$S \to 0 \qquad \text{as} \qquad T \to 0 \tag{7.4}$$

To see the necessity of this restriction, suppose that C did not vanish with T, but instead reached a constant value as the temperature approached zero. We could then remove C from the integral in (7.3) and obtain

$$S(T) = C \int_0^T \frac{dT}{T}$$

However, the integral is infinite![†] To put it another way, the entropy $S(T)$ does not exist if C tends toward a nonzero value as T approaches zero. In order for $S(T)$ to have a definite value the heat capacity must vanish with the temperature. Experiment confirms this restriction on C. For electrically insulating solids it is found that C is proportional to T^3 at low temperatures

$$C_{\text{ins}} = bT^3 \tag{7.5}$$

[†] Thus, $\int_{T_0}^{T} dT/T = \ln(T/T_0) \to \infty$ as $T_0 \to 0$.

For electrically conducting solids, which have electrons free to roam about, the low-temperature behavior of C has the form

$$C_{\text{cond}} = aT + bT^3 \tag{7.6}$$

The aT term is contributed by the conduction electrons. As the student can readily verify, both (7.5) and (7.6) lead to entropies which vanish as T tends to zero.

•••••• ••••••

7.2.1 (a) Using the heat capacities of (7.5) and (7.6) evaluate the low-temperature entropies of insulating and conducting solids. Verify that $S(T) \to 0$ as $T \to 0$.

(b) Show that the entropy vanishes at $T = 0$ provided

$$C \propto T^q$$

where q is any number greater than zero.

•••••• ••••••

Yet another confirmation of the third law is furnished by measurements of the melting pressure of ^{4}He. Under atmospheric pressure ^{4}He remains a liquid at the lowest temperature. However, if sufficient pressure is applied, the liquid solidifies. The melting curve marks the liquid–solid boundary on a pressure versus temperature plot. Figure 7.3 shows the behavior of the melting curve for ^{4}He. The slope of the melting curve is given by the Clausius–

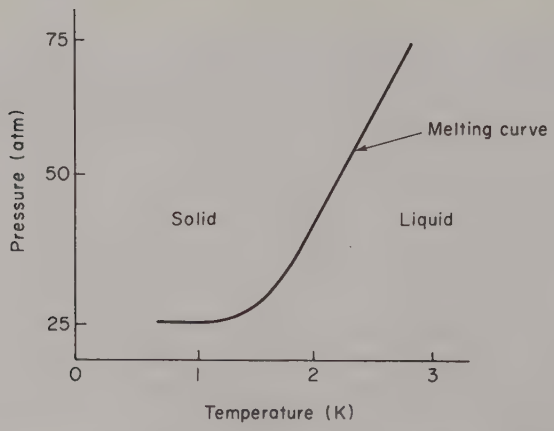


Figure 7.3. Melting pressure of ^{4}He. The slope of the curve dP/dT is given by the Clausius–Clapeyron equation (7.7). [After K. Mendelssohn, "Cryophysics." Wiley (Interscience), New York (1960).]

Clapeyron equation

$$\frac{dP}{dT} = \frac{L_f}{T(V_l - V_s)} \tag{5.9a}$$

where L_f is the latent heat of fusion at temperature T and $V_l - V_s$ is the volume change in the solid to liquid transition. Because melting is an isothermal process, L_f/T is just the entropy change of melting

$$S_l - S_s = \frac{L_f}{T} \tag{6.10}$$

The slope of the melting curve is therefore

$$\frac{dP}{dT} = \frac{S_l - S_s}{V_l - V_s} \tag{7.7}$$

Both S_l and S_s tend to zero as T is lowered toward 0 K. It follows that

$$S_l - S_s \to 0 \quad \text{as} \quad T \to 0$$

However, $V_l - V_s$, the volume change which accompanies melting, remains nonzero to the lowest temperature. It then follows from (7.7) that the slope of the melting curve must become zero near absolute zero. Figure 7.3 shows a striking confirmation of this prediction, the melting pressure being nearly constant below 1 K.

7.3 EXCEPTIONS AND RESTATEMENT

Despite an impressive string of successes, the unqualified statement of the third law is not universally valid. There are systems for which the entropy does not tend toward zero at low temperatures and which force us to restate the third law. The exceptions may be traced to irreversible processes which, instead of driving the system to equilibrium, leave it stranded in a nonequilibrium state.

The resulting difficulties are not heaped solely on the third law. Strictly speaking, the first and second laws no longer apply either. The reason for this is that thermodynamic properties, entropy, internal energy, and so on, are operationally defined only for equilibrium states. Thermodynamics therefore can make no positive statements about transformations which begin or end in nonequilibrium states.

To understand why these exceptions are not disastrous for the third law we must appreciate yet another aspect of entropy. The entropy and energy of a system are alike in one important respect. Both are sums of contributions from

various sources. For example, consider a gas of diatomic molecules. In addition to the familiar translational kinetic energy molecules can possess rotational and vibrational energies, all of which contribute to the internal energy of the gas. Now, the parallel between energy and entropy is this. Each and every mechanism that donates to the energy of a system also contributes to its entropy. The internal energy of a diatomic gas is a sum of the energies of rotation, vibration, translation, etc. Similarly, the entropy is the sum of a translational entropy, a vibrational entropy, and so on. This one-to-one correspondence of entropy and energy should not seem mysterious or miraculous. In a reversible process in which no work is performed the first law reduces to

$$dQ = dU$$

The entropy change is

$$dS = \frac{dQ}{T} = \frac{dU}{T}$$

It is immediately clear that if U is expressed as a sum of terms, then S may be expressed by a similar sum. The multiplicity of entropy and energy contributions is described by saying that a system possesses different *degrees of freedom*. In a well-behaved system, lowering the temperature reduces the energy and entropy associated with each degree of freedom. In particular, each contribution to the entropy tends toward zero with the temperature.

In the exceptional cases, where the system is in a nonequilibrium state, not every degree of freedom is involved. Generally, only the entropy contributions of a few degrees of freedom fail to drop toward zero. We say their entropy is "frozen in," an appropriate terminology in view of the temperature of the surroundings. The other degrees of freedom behave normally—their entropy contributions vanish at absolute zero. It is helpful to think of a system as a collection of subsystems. Each degree of freedom is a subsystem—a part of the whole. The exceptions to the unqualified form of the third law should be regarded as situations wherein the various subsystems are not all in mutual equilibrium. Those subsystems which are in mutual equilibrium obey the unqualified form of the third law. Their entropy contributions vanish at absolute zero. The rebellious subsystems are not in equilibrium. They defy a thermodynamic description. The subsystem concept lets us restate the third law in a universally valid form. Our earlier unqualified statement (The entropy of a system approaches zero as its temperature tends toward absolute zero.) is replaced by the following restatement, essentially as stated by Simon[†]:

[†] F. E. Simon, "Year Book of the Physical Society," p. 1. The Physical Society, London (1956). Adapted with permission of The Institute of Physics and The Physical Society.

> The contribution to the entropy of a system from each subsystem which is in internal thermodynamic equilibrium disappears at the absolute zero.

The Simon formulation of the third law has much to recommend it. It advises that we may apply the unqualified form of the third law to those subsystems (degrees of freedom) which are in internal (mutual) equilibrium. The sense of our original unqualified form is thereby maintained, and lets us salvage many results based upon it.

Exceptions to the unqualified form of the third law are to be found in the behavior of certain supercooled liquids. With care and patience it is possible to cool certain liquids below their customary melting point. Glycerol is an outstanding example for it can exist as a supercooled liquid down to the lowest temperatures, even though its normal melting point is about 19 °C. Even though the supercooled liquid is in a nonequilibrium state it is possible to measure its heat capacity. The specific heats of crystalline and supercooled liquid glycerol differ significantly from the melting point down to about 180 K. Below 180 K the heat capacities are nearly equal and both tend to zero at absolute zero. The difference in the two heat capacities may be traced to the translational degrees of freedom. In the liquid, the molecules are relatively free to roam about. In the crystalline solid the molecules form a lattice, thereby curtailing translational motion. The extra degrees of freedom give the supercooled liquid greater energy and entropy than the solid.

It is instructive to contrast the behavior of glycerol with that of a normal substance—one which obeys the unqualified form of the third law. To draw close parallels we must choose a normal substance which can exist in more than one form down to the lowest temperatures. Tin, sulfur, phosphine (PH_3), and cyclohexanol ($C_6H_{12}O$) are four such substances. For example, there are two forms of tin which, though chemically identical, have different crystalline structures. Grey tin is the stable form at low temperatures and white tin is the stable form at high temperatures. The transition temperature is 19 °C (= 292 K). Grey tin at 19 °C gradually (in a few hours) changes into white tin, releasing a heat of transition of 535 cal/mole.

In practice one tests whether or not the unqualified version of the third law is followed by comparing two determinations of the entropy difference. One determination follows from the heat of transition. A second evaluation is made using the measured heat capacities. We now outline this procedure and then illustrate with the data for tin and glycerol.

If form 1 undergoes a reversible transformation to form 2 at a temperature T, releasing a heat of transition $L_{1\to 2}$, the entropy change is

$$S_2(T) - S_1(T) = \frac{L_{1\to 2}}{T} \tag{7.8}$$

Next, we integrate $dS_2 = C_2\,dT/T$ from 0 to T, obtaining

$$S_2(T) - S_2(0) = \int_0^T C_2 \frac{dT}{T}$$

Combining this with the corresponding equation for form 1 gives a second expression for $S_2(T) - S_1(T)$,

$$S_2(T) - S_1(T) = S_2(0) - S_1(0) + \int_0^T (C_2 - C_1)\frac{dT}{T} \tag{7.9}$$

Comparing (7.8) and (7.9) shows that the entropy difference at absolute zero is

$$S_2(0) - S_1(0) = \frac{L_{1\to 2}}{T} - \int_0^T \frac{(C_2 - C_1)}{T}\,dT \tag{7.10}$$

The quantities on the right side of (7.10) are subject to measurement. If the right side is zero (to within the experimental uncertainty of about 0.1 eu/mole), we conclude that the substance obeys the unqualified form of the third law. Otherwise, we conclude that there is a frozen-in entropy.

For the grey-to-white tin transition

$$\frac{L_{g\to w}}{T} = \frac{535}{292}\ \frac{\text{eu}}{\text{mole}} = 1.83\ \frac{\text{eu}}{\text{mole}}$$

Measurements of the heat capacities give

$$\int_0^{292} \frac{(C_w - C_g)}{T}\,dT = 1.77\ \frac{\text{eu}}{\text{mole}}$$

These results establish that $S_w(0)$ and $S_g(0)$ are equal to within the experimental uncertainty and that tin obeys the unqualified form of the third law.

By contrast the data for crystalline and supercooled glycerol give[†]

$$\frac{L_{s\to l}}{T} \simeq 15\ \frac{\text{eu}}{\text{mole}}, \qquad \int_0^{292} \frac{(C_l - C_s)\,dT}{T} \simeq 10\ \frac{\text{eu}}{\text{mole}}$$

Thus

$$S_l(0) - S_s(0) \simeq 5\ \frac{\text{eu}}{\text{mole}}$$

The entropy difference here is well beyond the experimental uncertainties—the entropies of the supercooled liquid and the crystalline solid differ at absolute zero. The entropies differ because the liquid is more disordered than the crystalline solid phase. The supercooled liquid preserves the excess disorder—it appears experimentally as a frozen-in entropy difference.

[†] It is coincidental that the transition temperature is 292 K for both tin and glycerol.

Those exceptional substances whose entropy does not approach zero at absolute zero do not permit the ultimate low temperature to be attained. Their entropy is frozen-in. It is associated with degrees of freedom which are beyond external control. Only those degrees of freedom which remain in mutual equilibrium can be manipulated at the macroscopic level, and their contributions to the entropy vanish at absolute zero.

....

7.3.1 The molar heat capacities of two forms of a (hypothetical) substance are given by

$$C_\alpha = 0.09T + 0.003T^3 \quad \text{eu}$$

$$C_\beta = 0.01T \quad \text{eu}$$

The β-form undergoes a phase transition to the α-form at a temperature of 20 K. The measured heat of transition is $L_{\beta\to\alpha} = 190$ cal/mole. Determine whether or not the substance obeys the unqualified version of the third law, that is, whether both the α and β forms have the same entropy at absolute zero. Assume that the data for $L_{\beta\to\alpha}$ and the heat capacities are reliable to within $\pm 2\%$.

....

7.4 THE UNATTAINABILITY OF ABSOLUTE ZERO

Our earliest statement of the third law was simply that "You cannot reach absolute zero." The unattainability of absolute zero is a consequence of the fact that the entropy tends to zero with the temperature. In layman's language, it is a fatal attack of diminishing returns which prevents any scheme from lowering the temperature to absolute zero. While a general proof of unattainability can be formulated we shall be content to illustrate the idea with a particular example.

The illustration has been constructed so as to be formally identical to the following riddle: A man sets out to drive from Pittsburgh to New York City. He drives half way during the first day. On each succeeding day he drives half of the distance which remained at the end of the previous day. How many days are required to complete the journey? The answer is that the trip is never completed. So long as the distance traveled each day is a fraction of the remaining distance the trip can never be completed. An analogous fate befalls methods designed to lower temperatures to absolute zero. These schemes invariably lead to temperature changes proportional to the temperature itself—they lower the temperature by some fraction of the remaining "distance" to absolute zero.

It is not difficult to appreciate that any optimized method of lowering temperatures includes adiabatic and reversible transformations. Thus, if the system performs a positive amount of work adiabatically it does so at the expense of its internal energy—and a decrease in U lowers the temperature. It is also desirable to employ reversible processes. Irreversible processes generate entropy. But we know that entropy and temperature tend toward zero together. Entropy production via irreversible processes therefore opposes temperature reduction.

For our illustration we suppose that the entropy is directly proportional to the product of temperature and volume

$$S = aTV, \qquad a = \text{constant}$$

With the system initially in a state described by $T = T_0$, $V = V_0$, we perform a series of two-step processes, as indicated in Figure 7.4. The two straight lines on the S–T diagram are plots of $S = aTV$ for $V = V_0$ and $V = \frac{1}{2}V_0$. First, the system is isothermally compressed from V_0 to $\frac{1}{2}V_0$. This decreases the entropy, as indicated in the figure. The compression is followed by an adiabatic (and *isentropic*) expansion which returns the system to its initial volume. Since

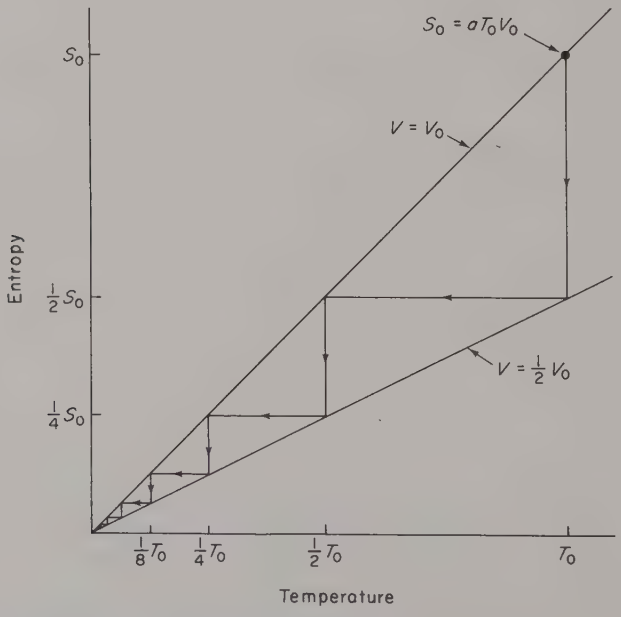


Figure 7.4. Isothermal compression reduces the volume and entropy. Isentropic expansion lowers the temperature. Because the temperature reduction is a fraction of the starting temperature no finite number of repetitions can reach absolute zero.

$S = aTV$ remains constant during the isentropic process, the volume change from $\frac{1}{2}V_0$ to V_0 must be accompanied by a temperature decrease from T_0 to $\frac{1}{2}T_0$. Overall, the two-step process returns the system to its initial volume and cuts its temperature in half. By repeating the two-step sequence the temperature can be halved again. And so it might go—*ad infinitum*. Each repetition halves the existing temperature. However, just as the man who forever goes halfway never reaches New York City, so the temperature never reaches absolute zero.

The third law places no restrictions on how closely one may approach absolute zero. A recently developed "dilution" refrigerator, which utilizes a mixture of ^{3}He and ^{4}He permits experiments to be performed in the millidegree (10^{-3} K) range. Details of dilution refrigerators are presented in Chapter 18.

REFERENCES

The "inside story" of the third law is told by:

F. Simon, The Third Law of Thermodynamics, "Year Book of the Physical Society," pp. 1–21. The Physical Society, London (1956).

Among the historical tidbits included is the revelation that Nernst never thought in terms of entropy!

A thorough discussion of all aspects of the third law is given in:

J. Wilks, "The Third Law of Thermodynamics." Oxford Univ. Press, London and New York (1961).

Some portions make use of elementary quantum mechanics. Nevertheless, the beginner will find helpful insights throughout the book.

A general proof of the unattainability of absolute zero can be found in:

M. W. Zemansky, "Heat and Thermodynamics," 5th ed., p. 497. McGraw-Hill, New York (1968).

Chapter 8

Thermodynamic Potentials

In the preceding chapter we completed our program of converting the first law into the mathematically viable form

$$dU = T\,dS - P\,dV \tag{8.1}$$

The internal energy U has proved to be a most useful thermodynamic property. Conceptually, we picture U as the sum of the kinetic and potential energies of the component parts of the system. Despite its conceptual clarity, the internal energy is ill-suited for the analysis of many thermodynamic processes. In this chapter we introduce three additional quantities, related directly to U, which find wide application in thermodynamics. These are the enthalpy H, the Gibbs free energy G, and the Helmholtz free energy F. Along with U, these functions are referred to as thermodynamic potentials. Each has the dimensions of energy. The potentials give thermodynamic theory greater flexibility and coherence. They also provide more direct links with experiment than could be managed using only the internal energy.

A major shortcoming of our formulation of thermodynamics is removed by introducing a fifth potential. In the interest of clarity we have confined our attention almost exclusively to closed systems consisting of a single chemical species. In practice it often becomes necessary to deal with open systems, which exchange matter with their surroundings, and with mixtures whose components undergo chemical or nuclear reactions. Both situations force us to deal with systems having a variable number of particles. The thermodynamic

parameter introduced to shoulder this task is called the chemical potential μ. The importance of the chemical potential is further enlarged by the fact that it is used extensively in statistical mechanics.

In the first section we present a brief introduction to the thermodynamic potentials. Later sections offer more leisurely and interpretive discussions of the potentials.

8.1 INTRODUCTION

The differential form of the first law (8.1) suggests that the entropy S and volume V be regarded as independent variables and that U be treated as a function of S and V

$$U = U(S,V)$$

If this is done, temperature and pressure become dependent variables whose relation to U, S, and V is readily determined. As a purely mathematical consequence of the fact that U is a function of the independent variables S and V we can write

$$dU = \left(\frac{\partial U}{\partial S}\right)_V dS + \left(\frac{\partial U}{\partial V}\right)_S dV \tag{8.2}$$

Comparing this with the first law (8.1) we see that[†]

$$T = \left(\frac{\partial U}{\partial S}\right)_V, \qquad P = -\left(\frac{\partial U}{\partial V}\right)_S \tag{8.3}$$

Entropy and volume are not necessarily the most suitable variables, a fact abundantly clear to anyone who has rummaged about a laboratory in search of an entropy meter. Let us therefore define a new thermodynamic potential in which entropy is relegated to the role of a dependent variable. This can be accomplished by adding and subtracting $S\,dT$ on the right side of (8.1). We then obtain

$$dU = T\,dS + S\,dT - S\,dT - P\,dV$$

The first pair of terms on the right side form a complete differential

$$T\,dS + S\,dT = d(TS)$$

[†] The equating of coefficients in (8.1) and (8.2) is mandatory because S and V are independent variables.

Transferring $d(TS)$ to the left side leaves

$$d(U-TS) = -S\,dT - P\,dV \tag{8.4}$$

The combination $U-TS$ is known as the Helmholtz free energy

$$F = U - TS \tag{8.5}$$

Equation (8.4) shows that F is a function whose differential involves dT and dV, but not dS

$$dF = -S\,dT - P\,dV \tag{8.5a}$$

Therefore F is to be regarded as a function of the independent variables T and V. The entropy and pressure become dependent variables. To see how S and P are determined we note that as a purely mathematical consequence of its dependence on T and V we can write

$$dF = \left(\frac{\partial F}{\partial T}\right)_V dT + \left(\frac{\partial F}{\partial V}\right)_T dV \tag{8.6}$$

Comparing this with (8.5a) shows that S and P are determined from

$$S = -\left(\frac{\partial F}{\partial T}\right)_V, \qquad P = -\left(\frac{\partial F}{\partial V}\right)_T \tag{8.7}$$

The Helmholtz free energy is of great importance in statistical mechanics. Once $F(T,V)$ has been obtained, the relation

$$P = -\left(\frac{\partial F}{\partial V}\right)_T$$

yields the equation of state of the system.

....

8.1.1 The free energy of a gas is given by

$$F = -RT\ln(V-b) - \frac{a}{V} + CT$$

where R, b, a, and C are constants. Show that the gas obeys the van der Waals equation of state

$$\left(P+\frac{a}{V^2}\right)(V-b) = RT$$

....

If we choose to eliminate V rather than S, we are led to the enthalpy function. Starting with the first law, $dU = T\,dS - P\,dV$, we play "complete the differ-

ential," this time by subtracting and adding $V\,dP$.† This gives

$$dU = T\,dS + V\,dP - P\,dV - V\,dP$$

The last two terms form a complete differential so that

$$d(U+PV) = T\,dS + V\,dP$$

The enthalpy H is defined as

$$H = U + PV \tag{8.8}$$

whence,

$$dH = T\,dS + V\,dP \tag{8.9}$$

and H emerges as a function of S and P. The fact that H may be regarded as a function of S and P leads to the results

$$T = \left(\frac{\partial H}{\partial S}\right)_P, \quad V = \left(\frac{\partial H}{\partial P}\right)_S \tag{8.10}$$

....

8.1.2 Using the fact that the enthalpy may be regarded as a function of the two variables S and P, obtain (8.10).

....

The final thermodynamic potential, the Gibbs free energy, may be generated from H or from F. Thus adding and subtracting $S\,dT$ on the right side of (8.9) leads to

$$d(H-TS) = -S\,dT + V\,dP$$

Alternately, adding and subtracting $V\,dP$ on the right side of (8.5a) results in

$$d(F+PV) = -S\,dT + V\,dP$$

Equivalent definitions of the Gibbs free energy are therefore

$$G = H - TS \tag{8.11}$$

and

$$G = F + PV$$

In terms of the internal energy,

$$G(T,P) = U + PV - TS \tag{8.12}$$

---

† Observe that an independent variable is converted to dependent status by completing the differential in which it stands.

The differential expressions above show that T and P are the natural choices for independent variables

$$dG = -S\,dT + V\,dP \tag{8.13}$$

Treating G as a function of T and P leads to the relations

$$S = -\left(\frac{\partial G}{\partial T}\right)_P, \qquad V = \left(\frac{\partial G}{\partial P}\right)_T \tag{8.14}$$

....

8.1.3 Using (8.13) and the fact that the Gibbs free energy is best regarded as a function of the independent variables T and P, derive (8.14).

8.1.4 The technical name for the game we call "complete the differential" is Legendre transformation. Suppose that the differential of a function has the form

$$d\Psi = u\,dx + v\,dy$$

In this form, x and y appear as the independent variables. It is desired to replace x by u as an independent variable.

(a) Show that this may be accomplished by completing the $u\,dx$ differential.
(b) Write down the differential of the function $\Omega(u, y) \equiv \Psi - ux$.
(c) Construct a function whose differential involves du and dv rather than dx and dy.

8.1.5 In advanced mechanics the Lagrangian L is regarded as a function of the variables x and v. Its differential therefore has the form

$$dL = f\,dx + p\,dv, \qquad f = \left(\frac{\partial L}{\partial x}\right)_v, \quad p = \left(\frac{\partial L}{\partial v}\right)_x$$

It is often desirable to deal with the Hamiltonian H, a function of p and x.

(a) By completing a differential, show that the expression $pv - L$ emerges as a function of p and x. (You will probably discover that $L - pv$ emerges somewhat more naturally, but $pv - L$ is the Hamiltonian!)
(b) The Lagrangian for a particle of mass m in a uniform gravitational field (gravitational acceleration g) is given by

$$L = \tfrac{1}{2}mv^2 - mgx$$

(v is the speed of the particle and x is its distance above some reference level of zero gravitational potential energy). Express the Hamiltonian as a function of p and x. Note that p is defined as $p = (\partial L/\partial v)_x$. What is the physical significance of p?

....

8.2 THE MAXWELL RELATIONS

The differentials of the four thermodynamic potentials are given by

$$dU = T\,dS - P\,dV \tag{8.1}$$

$$dF = -S\,dT - P\,dV \tag{8.5a}$$

$$dH = T\,dS + V\,dP \tag{8.9}$$

$$dG = -S\,dT + V\,dP \tag{8.13}$$

Equation (8.1) is the first law. The next three follow in straightforward fashion from our "complete the differential" program. Any number of mnemonics have been constructed to aid in remembering these equations. One such scheme uses the diagram of Figure 8.1. The four potentials label sides of a

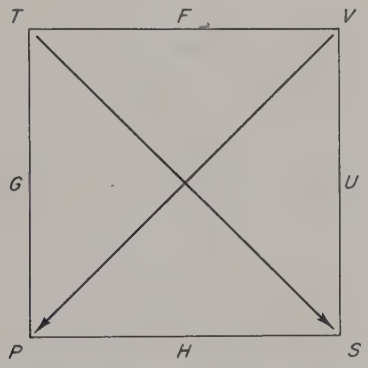
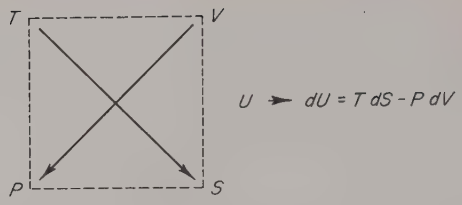


Figure 8.1. "The Friendly Vagabond Usually Shares His Precious Gifts."

square. The four variables S, T, P, and V label corners. A mnemonic phrase is used to remember the order of the symbols. As an example, there is the expression, "*T*he *F*riendly *V*agabond *U*sually *S*hares *H*is *P*recious *G*ifts." The student is invited to make up his own mnemonic. Note that each thermodynamic potential is flanked in the diagram by the two variables whose differentials appear in the expression for the differential of that potential. For example, U is flanked by S and V, and (8.1) shows that $dU = T\,dS - P\,dV$. The two arrows aid in the assignment of plus and minus signs. Part of a picture is worth at least 500 words.... As in the accompanying figure, going "with" an arrow rates a plus sign, opposing an arrow merits a minus sign. Test yourself!

$$dG = ?$$



By comparing two expressions for dU, (8.1) and (8.2), we arrived at (8.3),

$$T = \left(\frac{\partial U}{\partial S}\right)_V, \qquad P = -\left(\frac{\partial U}{\partial V}\right)_S \tag{8.3}$$

By cross differentiation we can obtain the first of the Maxwell relations. Thus, if we differentiate with respect to V on both sides of the first equation in (8.3), we obtain

$$\left(\frac{\partial T}{\partial V}\right)_S = \frac{\partial^2 U}{\partial V \partial S}$$

Similarly, applying $(-\partial/\partial S)_V$ to both sides of the second equation in (8.3) produces

$$-\left(\frac{\partial P}{\partial S}\right)_V = \frac{\partial^2 U}{\partial S \partial V}$$

The order of differentiation makes no difference

$$\frac{\partial^2 U}{\partial V \partial S} = \frac{\partial^2 U}{\partial S \partial V}$$

so that we obtain the result,

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V \tag{8.15}$$

This is one of the Maxwell relations. These relations are of considerable utility in thermodynamics. They permit the replacement of quantities which are difficult if not impossible to measure by quantities open to direct measurement. For example, $(\partial P/\partial S)_V$ is not directly measureable, whereas $(\partial T/\partial V)_S$ can be measured. Thus, if an insulated system undergoes a reversible process, its entropy remains constant ($dS = dQ/T = 0/T = 0$). If the change in volume ΔV and temperature ΔT are measured, we have

$$\frac{\Delta T}{\Delta V} = \left(\frac{\partial T}{\partial V}\right)_S$$

The remaining Maxwell relations follow by cross differentiation from (8.7), (8.10), and (8.14). They are

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V \tag{8.16}$$

$$\left(\frac{\partial T}{\partial P}\right)_S = \left(\frac{\partial V}{\partial S}\right)_P \tag{8.17}$$

$$-\left(\frac{\partial S}{\partial P}\right)_T = \left(\frac{\partial V}{\partial T}\right)_P \tag{8.18}$$

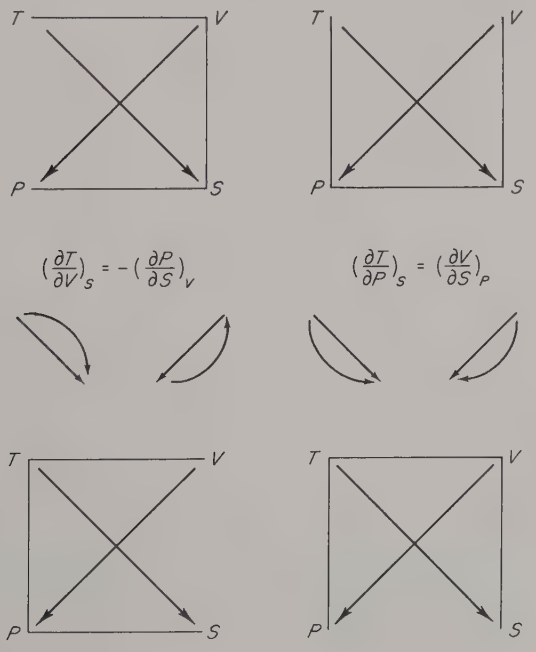
....

8.2.1 By performing cross differentiations, derive the three Maxwell relations (8.16)–(8.18).

....

The mnemonic diagram can be used to recall the Maxwell relations. Again, part of a picture is worth... Consider the accompanying figures. Test yourself! [Check your answers against (8.16) and (8.18).]

To illustrate the utility of the Maxwell relations, we show how (8.18) is used to derive an expression for $T\,dS$ needed later. If S is regarded as a function



of T and P,

$$dS = \left(\frac{\partial S}{\partial T}\right)_P dT + \left(\frac{\partial S}{\partial P}\right)_T dP \tag{8.19}$$

In a reversible process, $T\,dS$ equals the heat absorbed. If this heat exchange takes place at constant pressure, the heat capacity C_p enters as

$$(T\,dS)_P = dQ_P = C_p\,dT \tag{8.20}$$

From (8.19), with $dP = 0$,

$$(T\,dS)_P = T\left(\frac{\partial S}{\partial T}\right)_P dT \tag{8.21}$$

Comparing these two expressions for $(T\,dS)_P$ shows that

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_P \tag{8.22}$$

Inserting this result back into (8.19) yields

$$T\,dS = C_p\,dT + T\left(\frac{\partial S}{\partial P}\right)_T dP$$

If we now use the fourth Maxwell relation (8.18), we obtain the desired result

$$T\,dS = C_p\,dT - T\left(\frac{\partial V}{\partial T}\right)_P dP \tag{8.23}$$

By an argument similar to that leading to (8.23), and using (8.16), one can obtain another very useful expression for $T\,dS$. This task is left as an exercise for the student. The result is

$$T\,dS = C_v\,dT + T\left(\frac{\partial P}{\partial T}\right)_V dV \tag{8.24}$$

....

8.2.2 Use (8.19) and (8.23) to show that

$$\mu_{JT} \equiv \left(\frac{dT}{dP}\right)_H = \frac{1}{C_p}\left[T\left(\frac{\partial V}{\partial T}\right)_P - V\right]$$

where $(dT/dP)_H$ denotes the ratio of dT and dP for an isenthalpic process (H is constant) and μ_{JT} is called the Joule–Thomson coefficient. It measures the temperature change produced by a change in pressure under isenthalpic conditions. We encounter μ_{JT} again in the next chapter when we study the thermodynamics of gas liquefaction.

8.2.3 If the entropy is regarded as a function of T and V, its differential is

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV$$

Show that this form of dS leads to (8.24). You will need the Maxwell relation (8.16).

8.2.4 Show that (8.23) may be expressed in the form

$$T dS = C_p dT - \beta V T dP$$

where β is the coefficient of volume expansion.

....

8.3 HELMHOLTZ FREE ENERGY

In a purely mechanical system, where the concepts of heat and temperature may be disregarded, the adiabatic form of the first law ($Q = 0$) applies

$$W = -(U_f - U_i) = -\Delta U \tag{3.3'}$$

We interpret this by saying that the performance of work by the system is accomplished through a sacrifice of internal energy. Pressure, the thermodynamic variable which corresponds directly to the mechanical variable force, is related to the internal energy (8.3)

$$P = -\left(\frac{\partial U}{\partial V}\right)_S \tag{8.3'}$$

In a reversible adiabatic process the entropy remains constant, and we note that the derivative in (8.3) is evaluated holding S constant. Thus for strictly mechanical phenomena, the internal energy is the natural energy variable. The pressure, which embodies the dynamics of the system, follows from U via (8.3). However, thermodynamics flourishes because many processes are not adiabatic. Many important processes take place isothermally. In an isothermal process, the Helmholtz free energy $F = U - TS$ becomes the logical energy variable. To justify this statement we must show how the change in F is related to the work done in an isothermal process.

If T is held constant, we have from $F = U - TS$

$$\Delta F = \Delta U - T\Delta S$$

If the process is reversible $T\Delta S = Q$, the heat evolved. From the first law

$$-W = \Delta U - Q$$

we obtain

$$-\Delta F = W \qquad \text{(isothermal reversible process)} \tag{8.25}$$

We interpret this by saying that, in a reversible isothermal process, the work done by a system equals the decrease in its Helmholtz free energy. If the isothermal process is irreversible, we have $T\Delta S > Q$ and the same steps leading to (8.25) give instead $-\Delta F > W$. Thus, in general

$$-\Delta F \geqslant W \qquad \text{(isothermal process)} \tag{8.26}$$

The inequality in (8.26) stems from the fact that irreversible processes generate entropy over and above that needed to satisfy the first law. Contrasting (8.26) with the adiabatic form of the first law (3.3) reveals that the Helmholtz free energy is the natural energy variable in an isothermal process. This view is strengthened by (8.7)

$$P = -\left(\frac{\partial F}{\partial V}\right)_T \tag{8.7'}$$

This relation, which replaces (8.3), shows that the dynamics contained in P now follow most naturally from F rather than from U. This shift of importance from the internal energy to the Helmholtz free energy followed because our attention switched from adiabatic processes to isothermal processes.

Whereas (3.3) is an equality, the equality in (8.26) holds only for reversible processes. We interpret (8.26) by saying that work performed isothermally and reversibly is "paid for" by a decrease of free energy. For an irreversible process, only part of the decrease in free energy is transformed into work ($W < -\Delta F$). The remainder of the free energy decrease is degraded by the irreversible processes responsible for entropy production.

....

8.3.1 (a) Show that the Helmholtz free energy for 1 mole of an ideal gas can be expressed in the form

$$F = C_v T[1 - \ln T + (1-\gamma)\ln V - S_0]$$

where S_0 is a constant and $\gamma = 1 + R/C_v$.

(b) Verify the relation

$$P = -\left(\frac{\partial F}{\partial V}\right)_T$$

that is, show that it leads to the ideal gas equation of state.

(c) One mole of an ideal gas at temperature T undergoes a free expansion

which doubles its volume. Show that the free energy change is

$$\Delta F = -RT \ln 2$$

(d) Is the condition $W \leqslant -\Delta F$ satisfied? Can the process be identified as reversible or irreversible?

....

8.4 ENTHALPY AND THERMOCHEMISTRY

The four thermodynamic potentials were invented to make thermodynamics "easy." Each potential is the natural energy variable for certain classes of physical processes. The enthalpy finds perhaps its most important application in the field of thermochemistry.

Thermochemistry concerns the thermodynamic aspects of processes which convert heat into chemical energy, or vice versa. By chemical energy we mean the energy associated with the electrical forces which bind atoms together to form molecules. The basic idea is quite simple. One wants to learn how and how much energy is released in chemical reactions. Conveniently, much of the energy is released as heat. We can gain valuable insights into reaction energetics simply by measuring the heat evolved.

The most elementary aspects of thermochemistry involve only the first law. A central concept in thermochemistry is the heat of reaction. A chemical reaction may either liberate or absorb energy. If the reaction releases energy, the temperature of the materials in the reaction vessel increases, and heat is subsequently rejected to the surroundings. In this case the reaction is said to be *exothermic*. The converse, where heat is absorbed by the system, is referred to as an *endothermic* reaction. Very loosely, the heat of reaction is the heat transferred between the reacting system and its surroundings. In order to have an unambiguous definition of the heat of reaction it is necessary to refer the reactants and products to the same pressure and temperature. Thus the heat of reaction is defined as the heat absorbed or rejected in going from some reactant state of temperature T and pressure P to some final product state at the same temperature and pressure. Careful measurements of heats of reaction have aided in establishing tables of enthalpies, internal energies, and other thermodynamic potentials.

The majority of chemical reactions are conducted at a constant pressure of approximately 1 atm. If the reaction proceeds at constant pressure, the heat of reaction equals the enthalpy change. This follows from the first law of thermodynamics and the definition of enthalpy. Thus from the definition $H = U + PV$ we have, at constant pressure,

$$(dH)_P = dU + P\,dV \tag{8.27}$$

However, from the first law

$$dQ = dU + P\,dV$$

from which we conclude $(dH)_P = dQ$; the heat absorbed or rejected in a reversible isobaric process equals the enthalpy change. In a finite process,

$$Q = \Delta H \equiv H_f - H_i \tag{8.28}$$

This elementary observation is the foundation of thermochemistry. By measuring the heat of reaction we can determine the enthalpy change.

Experiment lets us measure enthalpy *changes*. There remains some freedom which is exploited by assigning a value of zero enthalpy to the elements when they are in the so-called standard state. The standard state of a substance is defined to be its most stable form at a pressure of 1 atm and a temperature of 25 °C. The enthalpy of each element in its standard state is zero—by definition. In principle the enthalpies of the elements in other states and of chemical compounds can be determined by measuring appropriate heats of reaction. (Enthalpy tables are frequently labeled enthalpy of formation or heat of formation.) Thus, we may conceive of forming 1 mole of sodium chloride by starting with 1 mole of metallic sodium and 1 mole of gaseous chlorine—each in its standard state—and permitting them to unite chemically. The initial enthalpy of the system is zero. If we took care to see that the sodium chloride ended up at a temperature of 25 °C and a pressure of 1 atm, the heat evolved would be approximately 98.23 kcal. This is the standard heat of formation of sodium chloride. The value of 98.23 kcal is not determined by carrying out the aforementioned reaction, but rather by indirect measurements. The standard heat of formation is generally indicated as $\Delta H_f°$. For example, the reaction equation

$$C + 2H_2 \rightarrow CH_4, \qquad \Delta H_f° = -17.88 \text{ kcal}$$

signifies that the standard heat of formation (denoted here by the subscript f) for methane is 17.88 kcal/mole. The negative sign reveals that heat is rejected in the reaction. In reactions other than those leading to the formation of a compound from its elements the heat of reaction is generally indicated by ΔH. For example, the oxidation of methane would be described by

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O, \qquad \Delta H = -192 \text{ kcal}$$

In all cases ΔH (or $\Delta H_f°$) is defined as the enthalpy of the products (pr) less that of the reactants (re)

$$\Delta H = H_{pr} - H_{re} \tag{8.29}$$

A little experience quickly convinces one that the chemical symbols appearing in the reaction equation can be regarded as enthalpies and treated algebraically.

Thus the reaction

$$AB + CD \rightarrow AC + BD$$

implies (assuming standard conditions for reactants and products)

$$\underbrace{(\Delta H_f^\circ)_{AB} + (\Delta H_f^\circ)_{CD}}_{H_{re}} = \underbrace{(\Delta H_f^\circ)_{AC} + (\Delta H_f^\circ)_{BD}}_{H_{pr}} - \Delta H$$

This is merely a definition of ΔH. Its utility will become evident shortly when we encounter it disguised as Hess's law.

The development of useful enthalpy tables relies on two facts. First, enthalpy is a state function, that is, a property of the thermodynamic state of a system which is independent of how the state is reached. One mole of sodium chloride in its standard state has an enthalpy of -98.23 kcal regardless of how it reached that state. This fact makes it possible to determine the enthalpies of certain compounds by indirect means and is the basis for Hess's law to be discussed shortly. The second fact is an empirical one, namely, that the enthalpy of a system depends strongly on temperature but only weakly on pressure. If we hold the temperature of a system constant and increase the pressure by 10 atm, the enthalpy change generally would be less than that produced by a 1 °C change in temperature. We can therefore often ignore pressure changes in enthalpy computations. A notable exception is encountered in gas liquefaction schemes where the slight pressure variation of enthalpy is crucial. We shall generally ignore the pressure dependence of enthalpy. This permits us to treat a chemical reaction as an isobaric process.

In a constant pressure reaction the heat rejected equals the enthalpy change

$$Q_p = H_{pr} - H_{re} \equiv \Delta H \tag{8.28'}$$

If the reactants are in their standard states, $H_{re} = 0$ and the heat rejected to the calorimeter equals the enthalpy of the products. For example, in the reaction

$$\tfrac{1}{2}H_2 + \tfrac{1}{2}Br_2 \rightarrow HBr$$

the heat rejected to the calorimeter is 8.66 kcal/mole of HBr formed. Thus $(\Delta H_f^\circ)_{HBr} = -8.66$ kcal/mole. Similarly, the complete oxidation of 1 mole of carbon releases 94.05 kcal, according to the reaction

$$C + O_2 \rightarrow CO_2$$

from which we conclude $(\Delta H_f^\circ)_{CO_2} = -94.05$ kcal/mole.

In many instances the enthalpy of formation must be determined indirectly. It is here that (8.28′) finds extensive application. By systematically combining observed heats of reaction it is possible to determine an enthalpy of formation not open to direct measurement. Various difficulties may rule out

such a direct measurement. For example, several reactions may proceed simultaneously. The measured heat of reaction is then a composite for the set of reactions. In other instances the reaction may proceed too slowly or too violently to permit a direct measurement of the heat of reaction. In still other situations the compound of interest may form only as an unstable intermediate which is subsequently transformed into a more stable compound. In such instances the enthalpy of formation must be determined by applying Hess's law, which states that the enthalpy change in a reaction is independent of the path by which the reaction proceeds. The validity of Hess's law is guaranteed by the fact that enthalpy is a state function. The enthalpy change depends on the initial and final states, but not on their connecting path. Algebraically, Hess's law is applied by adding or subtracting heats of reaction in such a way as to yield the unknown enthalpy. To illustrate the application of Hess's law, consider the problem of determining the enthalpy of formation of CO. A direct measurement of the heat of reaction for

$$C + \tfrac{1}{2}O_2 \rightarrow CO$$

is not possible because the carbon monoxide will be further oxidized to carbon dioxide. However, we can determine the enthalpy of formation of CO by regarding it as an intermediate step in the formation of CO_2, that is, we consider the reaction

$$C + O_2 \rightarrow CO_2, \qquad \Delta H_1 = -94.05 \text{ kcal}$$

as taking place according to the two-step sequence,

$$C + \tfrac{1}{2}O_2 \rightarrow CO, \qquad \Delta H_2 = ?$$

$$CO + \tfrac{1}{2}O_2 \rightarrow CO_2, \qquad \Delta H_3 = -67.64 \text{ kcal}$$

To apply Hess's law one merely replaces the chemical symbols by the corresponding enthalpies and recalls the definition $\Delta H = H_{pr} - H_{re}$ given by (8.29). For example,

$$CO + \tfrac{1}{2}O_2 \rightarrow CO_2$$

translates into

$$(\Delta H_f^\circ)_{CO} + \tfrac{1}{2}(\Delta H_f^\circ)_{O_2} = (\Delta H_f^\circ)_{CO_2} - \Delta H_3$$

The enthalpy of O_2 is zero so

$$(\Delta H_f^\circ)_{CO_2} - (\Delta H_f^\circ)_{CO} = \Delta H_3 = -67.64 \text{ kcal}$$

In similar fashion one finds from the enthalpy equations for the other two reactions,

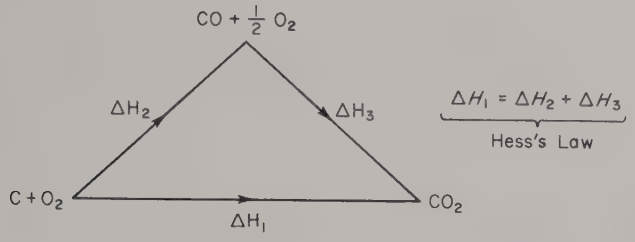
$$(\Delta H_f^\circ)_{CO_2} = \Delta H_1 = -94.05 \text{ kcal}$$

$$(\Delta H_f^\circ)_{CO} = \Delta H_2$$

From this set of equations and the measured values of ΔH_1 and ΔH_3 we find

$$(\Delta H_f^\circ)_{CO} = \Delta H_2 = \Delta H_1 - \Delta H_3 = -26.41 \text{ kcal}$$

These relationships may be expressed diagrammatically. The accompanying figure shows how Hess's law follows from the path independence of enthalpy change.



....

8.4.1 Determine ΔH for the reaction

$$C + H_2O \rightarrow CO + H_2$$

from the following information:

$$C + \tfrac{1}{2}O_2 \rightarrow CO, \qquad \Delta H = -26.4 \text{ kcal}$$

$$H_2 + \tfrac{1}{2}O_2 \rightarrow H_2O, \qquad \Delta H = -57.8 \text{ kcal}$$

8.4.2 Apply Hess's law to determine ΔH for the reaction

$$Mg + 2HCl \rightarrow MgCl_2 + H_2$$

from the measured heats of reaction for

$$Mg + Cl_2 \rightarrow MgCl_2, \qquad \Delta H = -153.2 \text{ kcal}$$

$$\tfrac{1}{2}H_2 + \tfrac{1}{2}Cl_2 \rightarrow HCl, \qquad \Delta H = -22.1 \text{ kcal}$$

....

8.5 GIBBS FREE ENERGY

The differential expression

$$dG = -S\,dT + V\,dP \tag{8.13}$$

establishes the Gibbs free energy as a function of the two independent variables P and T. The importance of G in chemistry and biology stems from the fact that many chemical reactions proceed at constant pressure and temperature.

If such reactions are reversible, (8.13) applies and shows that the Gibbs free energy is conserved. The Gibbs free energy is the natural energy variable in isothermal, isobaric processes. The quantity of signal importance in such a reaction is the change in G

$$\Delta G = G_{\text{pr}} - G_{\text{re}}$$

As we demonstrate below, a reaction will "go" spontaneously at constant temperature and pressure provided ΔG is negative, that is, if the system releases free energy to its surroundings. As a preliminary step then the chemist or biologist must ascertain whether or not a proposed reaction would result in a decrease in G. If ΔG should prove to be positive the reaction can still be made to go, but only through the addition of energy, perhaps in the form of heat. Such a change in G is not spontaneous and the pressure and/or temperature of the system are likely to be altered.

The decrease in G for a spontaneous isothermal and isobaric process is established as follows: From the definitions of F and G we find

$$G = F + PV$$

The change in G during an isobaric process is

$$\Delta G = \Delta F + P\Delta V$$

However, in an isobaric process

$$P\Delta V = P(V_{\text{f}} - V_{\text{i}})$$

equals the work done by the system, whence,

$$\Delta G = \Delta F + W \tag{8.30}$$

As we established in Section 8.3, the decrease in the Helmholtz free energy during an isothermal process is always greater than the work performed

$$-\Delta F \geqslant W \tag{8.26'}$$

or

$$\Delta F + W \leqslant 0$$

Comparing this with (8.30) establishes the desired result

$$\Delta G \leqslant 0 \tag{8.31}$$

the Gibbs free energy cannot increase during an isobaric and isothermal process. The equality in (8.31) corresponds to a reversible process, the inequality to an irreversible process. The spontaneous irreversible processes drive a system toward equilibrium, where its Gibbs free energy is a minimum.

An equilibrium state for a system at the same temperature and pressure as its surroundings is stable. Any spontaneous processes, which might upset the equilibrium, can occur only if they would decrease G. For a system already in the state of minimum G no such spontaneous processes can occur—the system is in equilibrium and the equilibrium is stable. Of course, equilibrium can be upset by external means, for example, by a change in the temperature of the surroundings. This causes the system to move toward a new equilibrium state. The free energy of the new state is again a minimum, but it is a minimum referring to the new temperature. An equilibrium value of G is a minimum compared to its values for other thermodynamic states of the same pressure and temperature.

Example 8.1

Two reactions of biological importance are the oxidation of glucose and photosynthesis. The overall result of the first of these may be represented by the scheme

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

The values of ΔG and ΔH, referred to the standard temperature of 298.16 K, are (per mole of glucose $C_6H_{12}O_6$)

$$\Delta G = -688 \text{ kcal/mole}, \qquad \Delta H = -673 \text{ kcal/mole}$$

Photosynthesis is equivalent to the reaction†

$$CO_2 + H_2O \rightarrow \{CH_2O\} + O_2$$

and has

$$\Delta G = 115 \text{ kcal/mole}, \qquad \Delta H = 112 \text{ kcal/mole}$$

For the oxidation of glucose, ΔG is negative and so the reaction proceeds spontaneously, liberating free energy. For photosynthesis, ΔG is positive and the reaction does not proceed spontaneously. The conversion of the radiant energy of sunlight into free energy drives the photosynthesis reactions. ✧

....

8.5.1 The two reactions of Example 8.1 result in entropy changes. Compute the entropy changes and explain why there is no violation of the second law. [*Hint:* Use (8.11) to relate ΔH and ΔG, noting that the reactions proceed isothermally.]

....

† The notation $\{CH_2O\}$ denotes the basic composition of a carbohydrate.

Another important class of processes which occur at constant pressure and temperature are changes of phase. Melting and vaporization are two familiar phase changes. Changes of phase are reversible processes, so that the Gibbs free energy is conserved, that is, G does not change. The subject of phase changes is important in many fields such as chemistry, physics, metallurgy, meteorology, and so on. Chapter 9 provides an introduction to the thermodynamics of phase changes.

In chemical and nuclear reactions, the number of particles of each reacting species can change. The thermodynamic potentials of the system generally change if the chemical composition is altered. In the next section we introduce the chemical potentials, quantities which describe the dependence of the thermodynamic potentials on the chemical composition.

8.6 CHEMICAL POTENTIAL

We have expressed the Gibbs free energy change by

$$dG = -S\,dT + V\,dP \tag{8.13}$$

This expression is incomplete under certain circumstances. For example, suppose we take a small pan of water as our system. We expose the water to the surrounding atmosphere on a warm day when the humidity is very low—the sort of day which makes you wish you did not have to study thermodynamics. If the water is at the same temperature and pressure as the surroundings, there will be no change in T or P. Under these conditions, (8.13) predicts that $dG = 0$, that is, that G is a constant. However, G is not constant! If we watch the system, the water in the pan, it disappears before our eyes—by evaporation. Thus G is an extensive quantity; it is directly proportional to the amount of water present. As water evaporates, G decreases. When the water has evaporated completely, $G = 0$. The deficiency of (8.13) is that it holds only for closed systems whose chemical composition does not change. This shortcoming is easily rectified. Consider an open system comprised of one phase of a single chemical species.[†] The pan of water is an example of such a system. Let N denote the number of molecules in the system, and note that N can change because the system is open. The Gibbs free energy depends on T, P, and N. If T, P, and N change by differential amounts, the change in G is

$$dG = \left(\frac{\partial G}{\partial T}\right)_{P,N} dT + \left(\frac{\partial G}{\partial P}\right)_{T,N} dP + \left(\frac{\partial G}{\partial N}\right)_{P,T} dN \tag{8.32}$$

[†] By a phase we mean matter which is homogeneous and which has definite boundaries.

The derivatives $(\partial G/\partial T)_{P,N}$ and $(\partial G/\partial P)_{T,N}$ are evaluated with N held constant. The constancy of N was implicit in the developments of Section 8.1 where we arrived at the results

$$\left(\frac{\partial G}{\partial T}\right)_P = -S, \qquad \left(\frac{\partial G}{\partial P}\right)_T = V$$

We now make this explicit by writing

$$\left(\frac{\partial G}{\partial T}\right)_{P,N} = -S, \qquad \left(\frac{\partial G}{\partial P}\right)_{T,N} = V \tag{8.33}$$

This converts (8.32) into

$$dG = -S\,dT + V\,dP + \left(\frac{\partial G}{\partial N}\right)_{T,P} dN \tag{8.34}$$

For a closed system, N is constant so that $dN = 0$ and (8.34) reduces to (8.13). The quantity

$$\left(\frac{\partial G}{\partial N}\right)_{T,P} \equiv \mu \tag{8.35}$$

is called the chemical potential; μ is a measure of how G is altered by a change in the number of particles present. With this definition, (8.34) becomes

$$dG = -S\,dT + V\,dP + \mu\,dN \tag{8.36}$$

It is important to note that the chemical potential is an intensive quantity. Thus μ is independent of N although it can vary with T and P. To see this, imagine that we double the size of the system—we mentally slap two of them together thereby making every extensive quantity twice as large without changing any of the intensive variables. The extensive quantities in (8.35) are S, V, dG, and dN. Each of these is doubled. In order for (8.35) to remain valid dT, dP, and μ must remain unchanged, that is, T, P, and μ must be intensive quantities. The fact that μ is an intensive variable means that it cannot depend on the amount of material in the system. Explicitly, μ cannot depend on N, although it will in general vary with T and P,

$$\mu = \mu(T, P), \qquad \left(\frac{\partial \mu}{\partial N}\right)_{T,P} = 0 \tag{8.37}$$

To account for the presence of several species of particles or more than one phase is routine from the bookkeeping point of view. For a system of C species we can write

$$G = G(T, P, N_1, N_2, \ldots, N_C)$$

where N_i denotes the number of molecules of the ith species. In place of (8.36) there then appears

$$dG = -S\,dT + V\,dP + \sum_i^{\mathrm{C}} \mu_i\,dN_i \tag{8.38}$$

where

$$\mu_i = \left(\frac{\partial G}{\partial N_i}\right)_{T,P} \tag{8.39}$$

is the chemical potential for the ith species. Each chemical potential is an intensive quantity, that is, μ_i can depend on T and P and the other μ's but not on the N_i, which are extensive variables. Suppose we consider a process which proceeds at constant pressure and temperature. Then $dT = 0$, $dP = 0$ in (8.38), and

$$dG = \sum_i \mu_i\,dN_i \tag{8.40}$$

We can "integrate" this equation to obtain

$$G = \sum \mu_i N_i \tag{8.41}$$

This integration is purely formal, and the result (8.41) follows directly from (8.40) upon remembering that both G and the N_i are extensive variables, whereas the μ_i are intensive.

We have seen how the introduction of the chemical potentials modified (8.13) by the addition of the term $\sum_i \mu_i\,dN_i$, giving (8.38). As G is related directly to the other thermodynamic potentials, they are modified in a similar manner. The expressions (8.1), (8.5a), and (8.9) are replaced by

$$dU = T\,dS - P\,dV + \sum_{i=1}^{\mathrm{C}} \mu_i\,dN_i \tag{8.42}$$

$$dF = -S\,dT - P\,dV + \sum_{i=1}^{\mathrm{C}} \mu_i\,dN_i \tag{8.43}$$

$$dH = T\,dS + V\,dP + \sum_{i=1}^{\mathrm{C}} \mu_i\,dN_i \tag{8.44}$$

From these expressions it is evident that μ_i can be defined not only by (8.39), but also as a derivative of U, F, or G,

$$\mu_i = \left(\frac{\partial U}{\partial N_i}\right)_{S,V} = \left(\frac{\partial F}{\partial N_i}\right)_{T,V} = \left(\frac{\partial H}{\partial N_i}\right)_{S,P} \tag{8.45}$$

With four equations featuring the chemical potentials there is now a choice of interpretations. The viewpoint adopted is likely to depend on the "ax one has to grind."

The biologist is likely to interpret μ_i from the standpoint of the Gibbs free energy. He is concerned with chemical reactions in which the temperature and pressure remain constant, and for which the Gibbs free energy is the natural energy variable.

The theoretical physicist or chemist concerned with the microscopic structure of matter is apt to interpret μ_i in terms of the Helmholtz free energy F. Indeed, we will adopt this viewpoint in our study of statistical mechanics.

The thermochemist and the low-temperature physicist who find it most convenient to deal with the enthalpy are prone to interpret μ_i in terms of their favorite thermodynamic potential.

There is even an interpretation of μ_i tailored to the needs of the student! Having struggled diligently to assimilate the concepts of heat and internal energy he now finds that the first law (with $T\,dS = dQ$) appears as

$$dU = dQ - P\,dV + \sum_{i=1}^{C} \mu_i\,dN_i \tag{8.42a}$$

This form permits an obvious interpretation of the chemical potentials. As its name suggests, the internal energy accounts for *all* forms of energy stored within the system. The right side of (8.42a) accounts for the different ways by which U can change. Thus the addition of heat dQ or the performance of mechanical work $-P\,dV$ can alter U, but changes in U also can arise through reactions which alter the composition of the system. It is precisely these changes in U which are reflected in the term $\sum_{i=1}^{C} \mu_i dN_i$ in (8.42a); the chemical potential simply reveals another of the many faces of energy.

....

8.6.1 An ideal monatomic gas has a molar heat capacity of $C_v = \frac{3}{2}R$. Recall that $k = R/N_0$ is the Boltzmann constant, and use the result of Problem 8.3.1 in conjunction with (8.45) to show that the chemical potential for an ideal monatomic gas may be expressed as

$$\mu = \frac{F}{N} = \tfrac{3}{2}kT[1 - \ln T - \tfrac{2}{3}\ln V - S_0]$$

....

Example 8.2 Phase Equilibrium

We wish to demonstrate that equilibrium between two (or more) phases of a substance requires equality of their chemical potentials. Thus, if a liquid and its vapor are to be in equilibrium it is necessary that $\mu_l = \mu_v$. Since phase transformations occur at constant temperature and pressure, (8.40) applies. We write (for two phases in equilibrium)

$$dG = \mu_1\,dN_1 + \mu_2\,dN_2 \tag{8.40a}$$

where 1 and 2 refer to the two phases. In Section 8.5 we established that G is a minimum for an equilibrium state. We envision a transformation in which a small amount of one phase converts to the other, while maintaining the equilibrium; for example, a drop of water vaporizes. On the one hand we must have $dG = 0$ to maintain G at a minimum, while on the other we must have

$$dN_1 = -dN_2 \tag{8.46}$$

to conserve mass. It then follows from (8.40a) that

$$\mu_1 = \mu_2 \tag{8.47}$$

Phase equilibrium requires equality of the chemical potentials. For multicomponent systems a generalization of (8.47) holds. The chemical potential of each species is the same for phases in joint equilibrium. For example, a pure substance at its triple point would have

$$\mu(\text{solid}) = \mu(\text{liquid}) = \mu(\text{gas}) \tag{8.47a}$$

A solution of bromine and water in equilibrium with their vapors would satisfy

$$\mu_{\text{Br}}(\text{liquid}) = \mu_{\text{Br}}(\text{gas}), \qquad \mu_{\text{H}_2\text{O}}(\text{liquid}) = \mu_{\text{H}_2\text{O}}(\text{gas}) \tag{8.47b}$$

Example 8.3 The Condition for Chemical Equilibrium

Chemical equilibrium is one of the conditions for thermodynamic equilibrium. In Section 8.5 we showed that the Gibbs free energy is a minimum for a state of thermodynamic equilibrium. We can use this fact to establish the general condition for chemical equilibrium.

For the purposes of illustration, consider the reaction whereby hydrogen and chlorine combine to form hydrogen chloride. This reaction may be represented in equation form by

$$\text{H}_2 + \text{Cl}_2 - 2\text{HCl} = 0$$

In general chemical reactions can be expressed in the form

$$\sum \nu_i A_i = 0 \tag{8.48}$$

in which the A_i denote chemical symbols and the ν_i are either positive or negative integers. For the hydrogen chloride reaction, $\nu_{\text{H}_2} = \nu_{\text{Cl}_2} = 1$ and $\nu_{\text{HCl}} = -2$.

If thermodynamic equilibrium prevails, the Gibbs free energy is a minimum. Suppose we alter the amounts of the various components under conditions of constant temperature and pressure. For processes in which only the N_i are

altered, (8.40) gives the change in G as

$$dG = \sum_{i=1}^{C} \mu_i \, dN_i \tag{8.40}$$

We want to apply this relation to a special situation, namely, where the N_i change in such a way as to maintain the existing chemical equilibrium. If G is a minimum at equilibrium, the changes in the N_i must be such that this minimum value is preserved, that is, we must have $dG = 0$ if the equilibrium is to persist

$$\sum_{i=1}^{C} \mu_i \, dN_i = 0 \tag{8.49}$$

The dN_i in (8.49) are related in a very special way to the ν_i in (8.48). To see just how these quantities are related, consider the hydrogen chloride reaction. Suppose we start by mixing 1 mole of hydrogen and 3 moles of chlorine. There will be many molecular reactions in which hydrogen chloride is produced. Eventually the reactions cease and we have chemical equilibrium. Now there is one very obvious way in which we can change the N_i without upsetting the existing equilibrium—and it is the only way. We can add, say, 1 mole of hydrogen and 1 mole of chlorine and remove 2 moles of hydrogen chloride. This does not unbalance the equilibrium because the added H_2 and Cl_2 combine to form the two moles of HCl which are removed. In general chemical equilibrium is maintained so long as

$$dN_i = K\nu_i \tag{8.50}$$

where K is an arbitrary constant, but must be the same for all reaction participants. In the example just cited, $K = N_0$. Inserting (8.50) into (8.49) leads to the desired relation—the condition necessary for chemical equilibrium

$$\sum_i \mu_i \nu_i = 0 \tag{8.51}$$

This equation illustrates quite clearly an earlier remark that the chemical potentials are not all independent variables. ✧

REFERENCES

The relationships between the second law and organic evolution are explored in:

H. F. Blum, "Time's Arrow and Evolution," 2nd ed. Harper, New York (1962).

Blum reveals the complementary aspects of the second law and Darwin's theory of natural selection. Note that Blum uses F to denote the Gibbs free energy G and E for the internal energy U.

Whether a thermodynamic system is open or closed may have unexpected consequences. Evidence supporting the "myth" that hot water cools faster than cold water is presented in:

G. S. Kell, Cooling of Hot and Cold Water, *Am. J. Phys.* **37**, 565 (1969).

An extensive treatment of thermochemistry will be found in:

G. W. Castellan, "Physical Chemistry." Addison-Wesley, Reading, Massachusetts (1964).

A more elementary approach is offered in:

B. H. Mahan, "Elementary Chemical Thermodynamics," Chapter 2. Benjamin, New York (1964).

Chapter 9

Phase Changes and Low-Temperature Physics

9.1 PHASE CHANGES AND PHASE DIAGRAMS

In the preceding chapter we defined a phase as matter which is homogeneous and which has definite boundaries. For example, the air in a balloon is in the gaseous phase and is separated from its surroundings by a definite boundary. Each chemically or physically distinct species present in a system qualifies as a component. The air is a single-phase multicomponent system, being a homogeneous mixture of nitrogen, oxygen, water vapor, and other gases. A container filled with water and water vapor constitutes a two-phase single-component system, there being a definite boundary between the liquid and gaseous phases as well as between the combined system and its surroundings.

A solid may be heated until its melts and becomes a liquid. The liquid in turn may be heated until it reaches its boiling point and vaporizes. Such phase changes are isobaric and isothermal processes. Heat is absorbed but the temperature does not rise until the phase change is complete. The bulk of the heat absorbed is used to alter the microscopic structure of the system. In melting for example, the atomic structure changes from the highly ordered array of the solid phase to the randomized liquid phase in which atoms behave like people in a panic-stricken crowd, jostling their neighbors as they move about erratically. In the melting process, atomic bonds must be broken. The heat of fusion, L_f, is the amount of heat absorbed in converting 1 gm of a

substance from solid to liquid. A large heat of fusion is indicative of strong atomic binding. Phase changes are reversible processes. When a liquid solidifies an amount of heat L_f is released.

....

9.1.1 Explain:

(a) why you may experience a chill upon emerging from a shower and
(b) how the air temperature may rise at the onset of a snowfall.

....

A heating curve illustrates the relationship between temperature and heat absorbed during a phase change. Figure 9.1 shows the heating curve for ice and water near the melting point. It plots temperature versus the amount of heat added Q. If the heat is added at a constant rate, Q is directly proportional to the heating time t. The heating curve shows that the sample starts as ice at $-5\,°C$. The addition of heat at a constant rate results in a uniform temperature rise until the ice reaches the melting point of $0\,°C$. No further increase of temperature results until all of the ice has melted. For a sample of m gm, the total heat absorbed during the phase change is mL_f, as indicated in the figure. Once all of the ice has melted, the addition of more heat causes the temperature rise to resume.

Example 9.1 A Latent Heat Measurement

If the heating rate is known, a graph of temperature versus heating time (that is, a heating curve) permits a determination of the latent heat. For example, suppose that heat is added at a constant rate of $\dot{Q} = 100$ W to

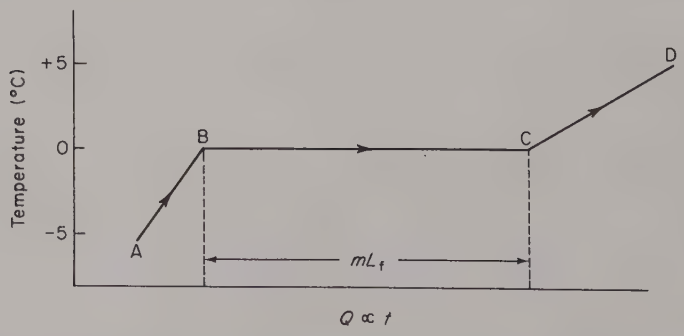


Figure 9.1. The heating curve for a sample of m gm of ice. The initial temperature is $-5\,°C$. Heat is added at a constant rate making Q proportional to t, the heating time. At B the ice has reached the melting point. During the change of phase (melting) the temperature remains constant ($0\,°C$). At C the ice has melted completely.

$m = 20$ gm of ice, and that the time during which the temperature remained constant (melting time) was 66 sec. The total heat absorbed during melting may be expressed as

$$mL_f = \dot{Q}t_{\text{melt}}$$

which gives

$$L_f = \frac{(100 \text{ joule/sec})(66 \text{ sec})}{(20 \text{ gm})} \frac{1 \text{ cal}}{4.18 \text{ joule}} = 79.1 \frac{\text{cal}}{\text{gm}} \diamondsuit$$

....

9.1.2 Explain how the slope of the heating curve is related to the specific heat of the sample. Is your explanation consistent with Figure 9.1, where the slope of the segment AB is approximately twice that of the segment CD? How do the specific heats of ice and liquid water compare?

....

A liquid at its boiling point may be vaporized by adding its heat of vaporization L_v. Part of the energy so supplied is used to free molecules from the surface of the liquid. A large portion of L_v is converted into work of expansion —the $P\,dV$ work done against the surrounding atmosphere.

Sublimation—the solid-to-vapor change of phase—is quite similar to vaporization, it being a process of boiling atoms from the surface of a solid rather than a liquid. The sublimation of dry ice (solid carbon dioxide) is a familiar example.

Example 9.2 The Method of Mixtures Again

A 50-gm ice cube at $-2\,^\circ$C is placed in 840 gm of water at 12 °C. The mixture is contained in an aluminum calorimeter with a mass of 200 gm and a specific heat of 0.215 cal/gm-°C. The initial temperature of the calorimeter is 12 °C. The specific heat of ice is 0.47 cal/gm-°C. The final temperature of the mixture is 7 °C. Using these data we wish to determine the heat of fusion of ice.

If we ignore all work of expansion, we can equate the heat rejected by the calorimeter and water with the heat absorbed by the ice cube. The 50 gm of ice is warmed to 0 °C. It then absorbs the heat of fusion and melts and, as water, is warmed to the final temperature of 7 °C. The heat absorbed by the 50 gm is

$$Q_{\text{abs}} = 50 \text{ gm}\left[\left(0.47 \frac{\text{cal}}{\text{gm-°C}}\right)(2\,^\circ\text{C}) + L_f + \left(1.00 \frac{\text{cal}}{\text{gm-°C}}\right)(7\,^\circ\text{C})\right]$$

The heat rejected by the water and calorimeter in cooling 5 °C is

$$Q_{\text{rej}} = 840 \text{ gm}\left(1.00 \frac{\text{cal}}{\text{gm-°C}}\right)(5\,^\circ\text{C}) + 200 \text{ gm}\left(0.215 \frac{\text{cal}}{\text{gm-°C}}\right)(5\,^\circ\text{C})$$

Equating Q_{abs} and Q_{rej} gives the heat of fusion as

$$L_f = 80.4 \text{ cal/gm}$$

The accepted value is 79.7 cal/gm suggesting that some form of heat or energy exchange has been ignored. In this type of experiment the heat absorbed or rejected by the thermometers, stirrers, and so on, must be included and the effects of radiation minimized or accounted for. ✧

•••••• ••••••

9.1.3 An immersion heater is placed in a mixture of 100 gm of ice and 150 gm of water (at 0 °C). The heater converts electrical energy into heat at a rate of 3000 W. Assume that all of the heat is transferred to the water-ice mixture. (This ignores the heat absorbed by the heater coil and container.) Determine the time required:

(a) to melt the ice,
(b) to bring the water to its normal boiling point,
(c) to completely evaporate the water.

•••••• ••••••

Melting and vaporization are termed first-order phase changes. The hallmark of a first-order phase transition is an entropy change. As a consequence of the entropy change there is an accompanying latent heat. For example, when ice melts the entropy of the resulting liquid exceeds that of the ice. The latent heat of fusion L_f is related to the entropy change ΔS and melting point T by

$$L_f = T\Delta S = T(S_l - S_s)$$

In the so-called higher-order phase changes the entropy is the same for both phases but other thermodynamic properties change. There is no latent heat.

•••••• ••••••

9.1.4 Starting from the thermodynamic relation

$$dS = \frac{dQ}{T}$$

derive the result

$$L_f = T\Delta S$$

Would the result be valid if a change of phase were not a reversible process?

•••••• ••••••

In Chapter 5 we indicated how the Clausius–Clapeyron equation could be used as a thermometer, or as the basis of a scheme for determining latent heats.

For one-component systems the Clausius–Clapeyron equation also may be used to map out important boundaries on what are known as *phase diagrams*. One-component systems such as the simple fluid or the paramagnetic solid are elementary in that their equations of state relate only three thermodynamic variables. For a fluid the equation of state may be stated as a relation giving the pressure in terms of the volume and temperature. Examples include our favorite—the ideal gas—

$$P = \frac{RT}{V}$$

and the van der Waals equation of state,

$$P = \frac{RT}{V-b} - \frac{a}{V^2}$$

A paramagnetic solid which obeys Curie's Law has an equation of state of the form (the volume V is regarded as a constant)

$$M(H,T) = CV\frac{H}{T}$$

Regardless of the functional form of the equation of state it can be represented geometrically as a surface so long as it relates just three variables. Figure 4.1 shows a segment of the P–V–T surface for the ideal gas. By relabeling the axes appropriately the same surface portrays the Curie paramagnet equation of state.

....

9.1.5 (a) Relabel the P–V–T axes of Figure 4.1 so that the surface represents the Curie paramagnet equation of state

$$M(H,T) = CV\frac{H}{T}$$

Note that this relabeling can be accomplished in more than one way. Of the three quantities P, V, and T, only one is an extensive variable. Likewise, of M, H, and T, just one is an extensive variable. Choose the relabeling scheme which places the extensive variables in correspondence. (You may wish to review Section 4.4.)

(b) Describe the three magnetic processes which correspond to the paths A→B→C→A in Figure 4.1.

....

The ideal gas, the van der Waals gas, and the Curie paramagnet share one common property. Unlike real substances they do not exhibit any changes

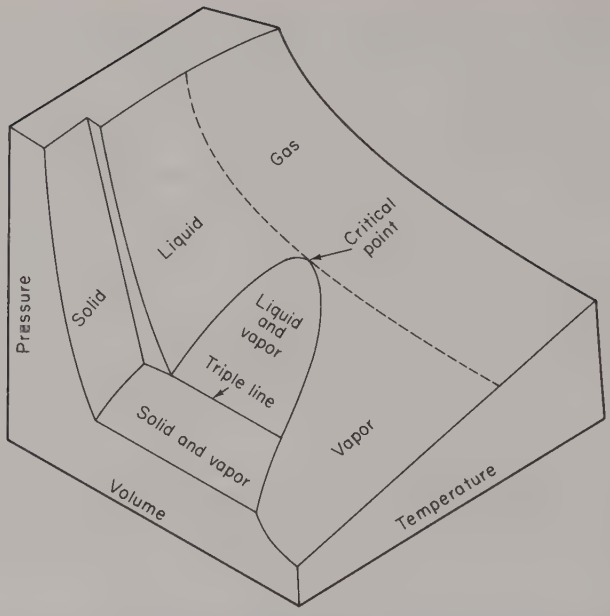


Figure 9.2. The *P–V–T* surface for carbon dioxide. The dashed line marks the critical isotherm. The appearance of the surface, when viewed along the volume axis, is shown in Figure 9.3.

of phase. Figure 9.2 shows the *P–V–T* surface of carbon dioxide (CO_2). The high-temperature low-pressure segments of the surface are very similar to that of the ideal gas. Various regions of the surface are labeled solid, liquid, and gas. These correspond to ranges of P, V, and T for which only a single phase can exist. Other portions of the surface locate states where two (or three) phases may coexist in a stable equilibrium. The line labeled "triple line" identifies a unique pressure and temperature and a range of volume (or density) over which all three phases can simultaneously exist. The point labeled "critical point" is of special significance. The critical point for a pure substance is located at a unique pressure, temperature, and density (P_c, T_c, ρ_c) which marks the upper terminus of the liquid–vapor coexistence region. The isotherm passing through the critical point is called the "critical isotherm." If a gas is compressed isothermally at a temperature in excess of the critical temperature (that is, along an isotherm above the critical point) it never undergoes condensation—it remains a gas. If it is compressed isothermally at a temperature less than T_c, a separation into liquid and vapor phases is observed. It is possible to start with a substance in the gaseous phase at a temperature $T_i < T_c$, raise its temperature above T_c, compress it isothermally,

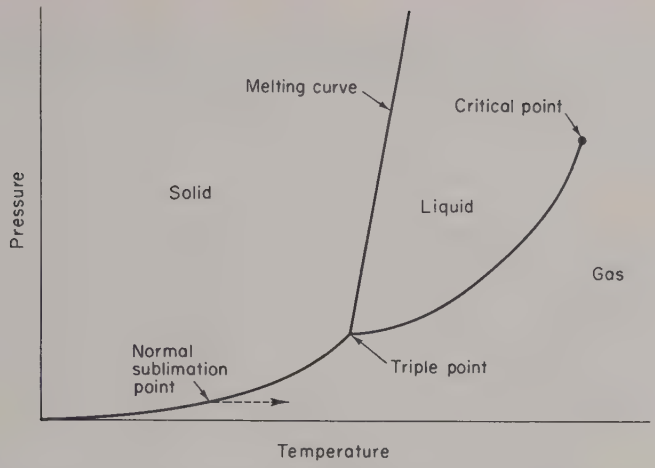


Figure 9.3. *P–T* diagram for carbon dioxide (not to scale). The lines define conditions for stable equilibrium between two phases. Their intersection is the triple point (5.1 atm, 216.6 K). The slopes of the three curves are determined by the Clausius–Clapeyron equation. Carbon dioxide is typical of most pure substances in that it expands upon melting. As a consequence the slope of its melting curve is positive ($dP/dT > 0$). The horizontal arrow at the normal sublimation point indicates that CO_2 sublimes at 1 atm (194.6 K). If the pressure is raised above 5.1 atm, heating the solid may result in melting. The values at the critical point are 72.9 atm and 309.1 K.

and then cool it to T_i, ending with a liquid. At no time would any condensation be evident. There simply would be a continuous transition from the gaseous to liquid phase.

If the *P–V–T* surface is projected onto the *P–T* plane we obtain a *P–T* phase diagram.† Figure 9.3 shows the *P–T* diagram for carbon dioxide. The triple line projects onto the *P–T* diagram as the triple point. The portion of the *P–V–T* surface labeled "liquid and vapor" appears as a line on the *P–T* diagram—the vaporization curve. The other coexistence regions likewise appear as lines on the *P–T* plot. The solid–liquid region of the surface projects onto the *P–T* plane as the line designated melting curve (also termed the "fusion line"). The solid–gas segment translates into the sublimation curve. The Clausius–Clapeyron equation determines the slopes of these three curves. The Clausius–Clapeyron equation was derived in Chapter 5

$$\frac{dP}{dT} = \frac{L_v}{T(V_g - V_l)} \tag{5.9a'}$$

† We shall use the explicit designation *P–T* diagram since there are many diagrammatic ways of representing the phases of a system.

where L_v is the latent heat of vaporization, the heat required to convert one mole of liquid to vapor and V_g and V_l denote the molar volumes of gas and liquid, respectively. A generalized form of (5.9a′), applicable to sublimation and melting as well as vaporization, is

$$\frac{dP}{dT} = \frac{L_{21}/T}{(V_2 - V_1)} \tag{9.1}$$

where L_{21} is the molar latent heat for the change from phase 1 to phase 2. The numerator L_{21}/T is the molar entropy change. In all three phase changes, melting, vaporization, and sublimation, the entropy increases. Stated differently, the associated latent heat is positive in all three cases. Vaporization and sublimation result in volume increases so that $V_2 - V_1$ is positive. It follows from (9.1) that dP/dT is positive for these two transitions

$$\left(\frac{dP}{dT}\right)_v > 0, \qquad \left(\frac{dP}{dT}\right)_{sub} > 0$$

Most substances, including carbon dioxide, expand upon melting, in which case

$$\left(\frac{dP}{dT}\right)_{melt} > 0, \qquad V_l > V_s$$

Although the form of the CO_2 diagram is typical its scale is not. The triple-point vapor pressure is unusually high—over 5 atm. As a consequence, the normal melting point lies well below the triple point and the solid sublimes—converts from solid to vapor—at a pressure of 1 atm. For this reason solid CO_2 has become better known as dry ice.

There are a few substances which contract upon melting. For these materials the Clausius–Clapeyron equation reveals that the melting curve has a negative slope

$$\left(\frac{dP}{dT}\right)_{melt} < 0, \qquad V_l < V_s$$

Water is among the substances which contract upon melting. Figure 9.4 shows the P–T diagram for water and reveals the negative slope of the melting curve. The dashed line in Figure 9.4 reveals that an increase in pressure can cause melting. Pressure-induced melting permits glaciers to flow around obstacles. The refreezing which follows the removal of pressure is called regelation. This phenomenon may be demonstrated using a block of ice, a large weight, and a fine wire. The wire is looped over the block of ice with the weight attached beneath. The pressure beneath the wire causes the ice to melt and the wire slowly falls through the block. The water refreezes behind the wire. Thus, even though the wire passes through it, the block ends up unchanged.

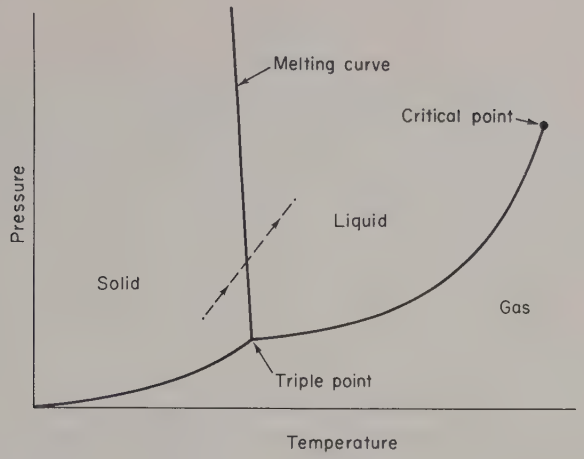


Figure 9.4. The P–T diagram for water (not to scale). Only a small portion of the P–T diagram is shown. Not indicated are the many high-pressure phases of ice. Because water expands upon freezing the slope of the melting curve is negative ($dP/dT < 0$). An increase in pressure (as suggested by the dashed line) can thereby induce melting. The values at the critical point are $P_c = 218.2$ atm and $T_c = 647.3$ K. The values at the triple point are $P_3 = 4.58$ Torr and $T_3 = 273.16$ K.

Example 9.3 The Heat of Sublimation of CO_2—An Estimate

In Chapter 5 we showed how the heat of vaporization could be determined by applying the Clausius–Clapeyron equation to vapor pressure data. Here we indicate how it may be used to estimate the heat of sublimation of solid carbon dioxide. If we have two data points on the sublimation curve, reasonably close together, we can determine L_{sub} by setting

$$\frac{dP}{dT} \simeq \frac{\Delta P}{\Delta T} = \frac{P_2 - P_1}{T_2 - T_1}$$

where (P_1, T_1) and (P_2, T_2) locate the two pressure–temperature data points. This approximation treats the sublimation curve as a straight line between the points (P_1, T_1) and (P_2, T_2). The latent heat of sublimation then follows from (9.1) as

$$L_{sub} = T(V_g - V_s)\frac{\Delta P}{\Delta T}$$

Since

$$V_g \gg V_s$$

we may safely ignore V_s by comparison to V_g. Furthermore, if we treat the

gas as ideal and use $PV_g = RT$, we obtain

$$L_{sub} \simeq \frac{RT^2}{P}\frac{\Delta P}{\Delta T}$$

This expression gives the molar heat of sublimation since we inserted the volume of 1 mole of gas.

Figure 9.3 can be used to make an estimate of L_{sub} for CO_2. The two points used are the normal sublimation point

$$P_1 = 1 \text{ atm}, \qquad T_1 = 194.6\,\text{K}$$

and the triple point

$$P_2 = 5.1 \text{ atm}, \qquad T_2 = 216.6\,\text{K}$$

For T and P we use the average values between the two points

$$P \simeq 3 \text{ atm}, \qquad T = 206\,\text{K}$$

The resulting value of L_{sub} is

$$L_{sub} \simeq 5.3 \text{ kcal/mole}$$

If we divide L_{sub} by the molecular weight of CO_2 (44 gm/mole), we obtain the specific heat of sublimation as 121 cal/gm. This is in fair agreement with the experimental value of 136.9 cal/gm. ✧

....

9.1.6 (a) Determine the entropy change when 1 mole of CO_2 sublimes.
(b) Is the majority of the heat of sublimation used to change the internal energy or to perform work of expansion? Use the first law to make a rough (order-of-magnitude) calculation to support your answer.

....

9.2 LIQUEFACTION AND CRYOGENICS

Cryogenics is the science of low temperatures. It is a mushrooming field which continually finds new and wider applications in research and technology. The history of cryogenics is delightfully recounted in Mendelssohn's popular work, "The Quest for Absolute Zero."† The early history of cryogenics is largely the story of pioneering efforts to liquefy gases. The liquefaction of oxygen by Cailletet and Pictet in 1877 was followed by a parade of successes. In succession, nitrogen, hydrogen, and finally helium were liquefied. The

† K. Mendelssohn, "The Quest for Absolute Zero." McGraw-Hill, New York (1966).

TABLE 9.1

Boiling Points and Heats of Vaporization of Cryogenic Fluids

Substance	Normal boiling point (K)	Heat of vaporization (cal/gm)
^4He	4.2	4.9
H_2	20.3	108.0
Ne	27.1	20.5
N_2	77.3	47.6
F_2	85.0	41.2
Ar	87.2	39.0
O_2	90.1	50.9
CH_4	111.6	121.9

normal boiling points of several cryogenic liquids are given in Table 9.1, along with their latent heats of vaporization. We will discuss only the basic physics underlying liquefaction schemes. Detailed descriptions of liquefiers and their operation will be found in the references.

There are two recognized techniques for liquefying gases. These are the external work and the internal work methods. Both methods act to reduce the kinetic energy of the molecules to a point where the attractive intermolecular forces cause the molecules to coalesce as liquid droplets. In the external work scheme a gas is expanded so as to perform work against its surroundings. If the expansion is adiabatic or nearly so, the work is done at the expense of the internal energy of the gas, and the temperature decreases. There are two primary drawbacks to the external work method. An expansion engine necessarily involves moving parts, the lubrication of which is difficult at low temperatures. Equally important is the fact that the temperature drop produced by an adiabatic expansion decreases in direct proportion to the initial temperature.

....

9.2.1 The equation describing the adiabatic expansion of an ideal gas is (see Example 4.2 and Problem 4.5.3)

$$\left(\frac{T}{T_0}\right)^{\gamma} = \left(\frac{P}{P_0}\right)^{\gamma-1}$$

Take $\gamma = 7/5$, a value appropriate for diatomic gases, and compute the final temperature for an adiabatic expansion which starts at $P = 100$ atm, $T = 300$ K, and ends at $P = 1$ atm. While real gases do not exhibit ideal behavior at 100 atm the resulting temperature is not unrealistic. How does your result compare with the normal boiling points (NBP) of oxygen and nitrogen?

....

The internal work method employs a throttling process, in which gas is forced through a narrow valve. To understand the meaning of the phrase internal work, we must anticipate two results of the kinetic theory of gases (Chapter 11). One result is that the temperature of a gas is a measure of the kinetic energy of the molecular motions. The second result is that, on the average, the intermolecular forces are attractive. During a throttle the average distance between molecules increases. Work is done by the molecules against the attractive intermolecular forces. The work increases the potential energies of the molecules at the expense of their kinetic energies. The decrease in kinetic energies registers as a temperature drop. The work is internal because the forces act between molecules in the gas, in contrast to the external work method where work is performed by forces pushing against the surroundings.

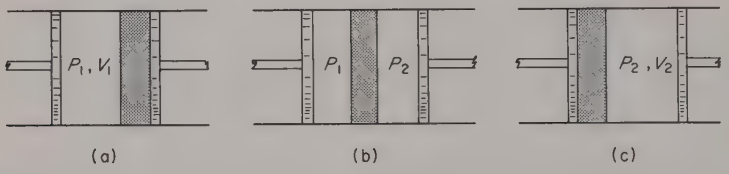


Figure 9.5. Three stages of the Joule–Thomson porous-plug experiment. Initially (a), the gas is to the left of the porous plug (dotted area) at a high pressure. The pressure forces the gas to seep through the pores of the plug. At an intermediate stage (b) a portion of the gas has wiggled through the plug. The piston on the right side expands, maintaining a constant pressure P_2. The left piston maintains a constant pressure P_1. The final state (c) finds all of the gas on the right side of the plug at the lower pressure P_2.

Thermodynamically, the throttling process is equivalent to the Joule–Thomson porous-plug experiment. In this experiment, gas is forced through a porous ceramic plug (see Figure 9.5). The gas starts on the left side of the plug at P_1, V_1 [part (a)]. As the piston forces gas through the plug, a second piston travels to the right. The pressure on the left side is maintained constant at P_1 throughout the process. The pressure on the right side is maintained constant at P_2. The final configuration [part (c)] has the gas on the right side of the plug at P_2, V_2. The entire system is insulated so that no heat is absorbed or rejected during the process. The total work done by the gas is (recall that work done isobarically is $P\Delta V$)

$$W = P_1(0-V_1) + P_2(V_2-0) = P_2V_2 - P_1V_1$$

With $Q = 0$, the first law gives

$$0 = \Delta U + W$$

Letting U_1 and U_2 denote the initial and final internal energies we find

$$U_1 + P_1V_1 = U_2 + P_2V_2 \tag{9.2}$$

However, $U+PV$ is the enthalpy H. We may therefore conclude that the enthalpy of the gas is left unchanged,

$$H_1 = H_2 \qquad \text{(porous-plug experiment or throttling process)} \quad (9.2a)$$

We should not jump to the conclusion that the throttling process is isenthalpic. It is not—because the process is irreversible. The intermediate states of the gas during the throttle are not equilibrium states. No infinitesimal change can reverse the gas flow through the porous plug (or throttle valve). The gas always flows from the high-pressure side to the low-pressure side. Nevertheless it is essential to realize that the enthalpy is left unchanged by the throttling process.

In a gas liquefaction apparatus the porous plug is replaced by a narrow valve. The gas pressure is perhaps 200 atm on one side, and only slightly above 1 atm on the other. If gases were ideal, there would be no temperature change upon throttling. This follows from the fact that the enthalpy of an ideal gas depends on the temperature alone. If there is no enthalpy change in an ideal gas, there can be no temperature change. In real gases, H depends on the pressure as well as on the temperature. The throttling of a real gas may result in cooling or heating, depending upon the value of the Joule–Thomson coefficient

$$\mu_{JT} = \left(\frac{\partial T}{\partial P}\right)_H$$

Experimentally μ_{JT} is determined as follows: An initial temperature and (high) pressure (T_1, P_1) are selected. The throttling process is then performed, carrying the gas to some chosen post-throttle pressure. The post-throttle temperature is measured. These data give a pair of T–P points which lie on the same isenthalp, that is, the two data points correspond to the same enthalpy. The process is then repeated, starting with the same initial T–P values throttling down to a different final pressure. In this manner a series of T–P points are obtained, all lying on the same isenthalp. The slope of this isenthalpic curve equals μ_{JT} for that value of the enthalpy. If the series of measurements is repeated for a different set of pre-throttle T–P values, a different isenthalp can be constructed, its slope giving the corresponding values of μ_{JT}. Figure 9.6 shows the typical form of such isenthalps. As the figure reveals, the Joule–Thomson coefficient is positive over a definite range of T–P values. This region is bounded by the dashed line known as the inversion curve on which μ_{JT} is zero. Cooling always results if the initial T–P point for the throttle lies within the region of nonnegative values of μ_{JT}. This conclusion follows most clearly from Figure 9.6 itself. It must be borne in mind that throttling always carries the gas to a lower pressure, that is, to the left on the T–P plot.

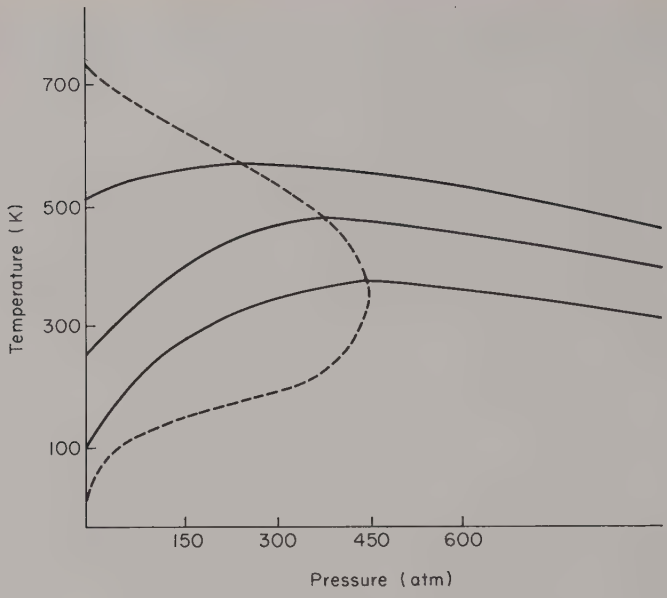


Figure 9.6. Typical form of isenthalps (solid lines) and inversion curve (dashed line, $\mu_{JT} = 0$). The throttle always carries the gas to the left, that is, to a lower pressure. If the throttle starts on the inversion curve, cooling results.

The essential elements of a Joule–Thomson liquefaction system are shown in Figure 9.7. First, the gas is compressed. This raises its temperature. The high-pressure gas is cooled and then throttled. When the system is operating in a steady state the throttling process causes a fraction of the gas to condense into the liquid phase. The remainder flows back to the compressor through a heat exchanger. The heat exchanger effects an internal exchange of heat. In the heat exchange operation, the pre-throttle gas is cooled by guiding it toward the J–T valve through coils bathed in a stream of cooler post-throttle gas which is thereby warmed as it returns to the compressor.

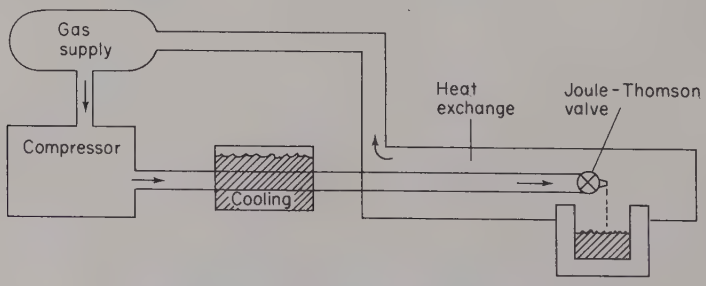


Figure 9.7. Schematic of Joule–Thomson liquefier. (See text for details.)

Low-temperature physics was an academic subject until the development of a commercial helium liqufier. The pioneering design was that of Kapitza in 1934. Kapitza used a combination of Joule–Thomson throttles and an expansion engine to liquefy helium. In 1946 Collins perfected a modification of Kapitza's liquefier. The Collins Machine turns out liquid helium on a commercial scale, providing the refrigerant for low-temperature research and technology.

Ultra-low temperatures have been achieved by using a magnetic version of the external work method. Magnetic cooling is discussed in the next section. A recent development which shows promise is the dilution refrigerator, which employs a mixture of liquid ^{3}He and ^{4}He. The ^{3}He–^{4}He refrigerator, whose operation is based on quantum mechanical effects, has made it possible to conduct experiments at temperatures near 10 millidegrees (0.01 K). Dilution refrigeration is discussed in Chapter 18.

9.3 MAGNETIC COOLING

The adiabatic expansion of a gas lowers its temperature. From a microscopic viewpoint the cooling stems from a conversion of molecular kinetic energy into external mechanical work. The adiabatic demagnetization of a paramagnetic substance also results in cooling. The energy converted into work is the potential energy of the magnetized material. Whereas an expanding gas sacrifices energy by pushing outward against atmospheric pressure, a paramagnetic substance lowers its internal energy by doing work against the currents and emf's which maintain its magnetic "atmosphere."

The energy decrease stemming from adiabatic demagnetization follows at once from the magnetic version of the first law of thermodynamics (4.16). In magnetic cooling experiments, the magnetic work $-H\,dM$ greatly exceeds the mechanical work $P\,dV$. Accordingly, we ignore the mechanical work term in (4.16) which then reads

$$dQ = dU - H\,dM \tag{4.16'}$$

where H is the applied magnetic field and M is the magnetic moment of the material. For an adiabatic process, $dQ = 0$, whence

$$dU = H\,dM \tag{9.3}$$

A paramagnetic material has a positive susceptibility χ

$$\frac{M}{V} = \chi H, \qquad \chi > 0 \tag{9.4}$$

It follows that a demagnetization—a decrease in H—diminishes M and thus U.

The reduction of the internal energy is revealed as a drop in temperature. The possibility of cooling by adiabatic demagnetization was suggested in 1926 by Debye and, independently, by Giauque, who also performed the first successful experiments.

The paramagnetic substance must have a comparatively large magnetic susceptibility. A number of hydrated salts containing paramagnetic ions have proved satisfactory in this respect. For example, in chromium potassium alum $K_2SO_4 \cdot Cr_2(SO_4)_3 \cdot 24H_2O$ the Cr^{3+} ions are the paramagnets which give rise to the large susceptibility. In another widely used salt, cerium magnesium nitrate, cerium ions are the magnetically active particles. Whereas an ordinary paramagnetic substance has a magnetic susceptibility of $\chi \sim 10^{-6} - 10^{-5}$ cgs units, the salts employed in magnetic coolers have $\chi \sim 10^{-2} - 1$ cgs units, over their operating temperature range of roughly 0.01–1 K.

The process of magnetic cooling starts with the salt pill suspended inside an evacuated container. The container in turn is cooled to near 1 K by bathing it in liquid helium. A strong magnetic field is applied and a small amount of gaseous helium is admitted to the sample container. The application of the magnetic field would increase the internal energy of the salt if the gaseous helium were not present. The gaseous helium conducts the added energy, as heat, to the surrounding liquid helium bath.† The magnetization is very nearly an isothermal process. This series of steps leaves the salt pill at the bath temperature (~ 1 K) in a strong magnetic field (10–50 kG). The gaseous helium is then pumped out of the sample container, leaving the paramagnetic sample thermally isolated from its surroundings. The magnetic field is then removed suddenly. The adiabatic demagnetization cools the salt. It is important to understand that the cooling is a result of an adiabatic process—the salt does not reject any heat to its surroundings. It is cooled because it performs magnetic work against its surroundings at the expense of its internal energy.

The thermodynamics of the cooling is best interpreted in terms of an entropy–temperature (S–T) diagram. Figure 9.8 indicates the behavior of the entropy as a function of the temperature for several different values of H, the applied magnetic field strength. At a fixed temperature the entropy decreases as H increases. This is in accord with the "entropy measures disorder" idea. The field acts to align the magnetic moments of the paramagnetic ions, an order- producing action. The stronger the magnetic field, the more complete is the alignment. In Figure 9.8, the directed arrows indicate the key steps in magnetic cooling. The isothermal magnetization $1 \rightarrow 2$ is followed by the adiabatic demagnetization $2 \rightarrow 3$. The adiabatic demagnetization is isentropic

† The gaseous helium serves only to conduct heat from the sample to the bath. It remains gaseous because its pressure is less than the saturated vapor pressure of helium at the bath temperature.

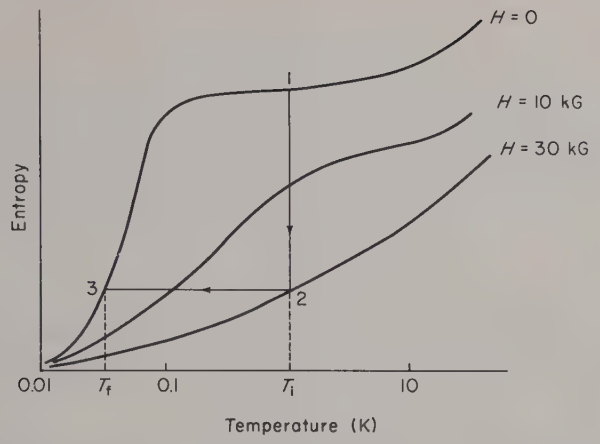


Figure 9.8. Magnetic cooling proceeds via the two-step process indicated. An isothermal magnetization at a temperature of 1 K carries the system to a state of lower entropy. The adiabatic demagnetization lowers the temperature to less than 0.1 K.

because it is (very nearly) a reversible process, that is, $dQ = T\,dS$ applies to a reversible process, and if $dQ = 0$ (adiabatic), then $dS = 0$ (isentropic). It is essential to start the demagnetization from as low a temperature as possible because the final temperature is directly proportional to the initial temperature.

....

9.3.1 (a) Assume that the internal energy U depends on T alone and show that the isothermal magnetization results in an entropy decrease given by

$$\Delta S \equiv S(H,T) - S(0,T) = -\frac{1}{T}\int_0^H H\,dM$$

(b) Show that the entropy decrease for a material obeying Curie's law ($\chi = C/T$) is

$$\Delta S = -\frac{CVH^2}{2T^2}$$

....

The thermodynamics of adiabatic demagnetization can be developed swiftly by exploiting analogies between magnetic and mechanical variables. In Chapter 4 we derived an expression for the thermodynamic work performed by a magnetic system

$$dW_{\text{mag}} = -H\,dM \qquad (4.15')$$

Adding to this the mechanical work $P\,dV$ gives the total thermodynamic work

$$dW = P\,dV - H\,dM \tag{9.5}$$

In Chapter 8 the magnetic work term was ignored. In the present situation it is quite legitimate to ignore the $P\,dV$ term in favor of the magnetic work $-H\,dM$. Observe that P and H are both intensive variables; V and M are both extensive variables. From (9.5) it follows that the analogs are

$$P \leftrightarrow H, \qquad V \leftrightarrow -M \tag{9.6}$$

A number of useful relations among the thermodynamic variables were derived in Chapter 8. Among these were the Maxwell relations. In particular, the derivatives of the entropy and volume were found to be related by

$$-\left(\frac{\partial S}{\partial P}\right)_T = \left(\frac{\partial V}{\partial T}\right)_P \tag{8.18}$$

Replacing P and V by their magnetic analogs results in

$$\left(\frac{\partial S}{\partial H}\right)_T = \left(\frac{\partial M}{\partial T}\right)_H \tag{9.7}$$

By regarding the entropy as a function of the independent variables T and H we are led to the magnetic version of (8.23). Thus,

$$dQ = T\,dS = T\left(\frac{\partial S}{\partial T}\right)_H dT + T\left(\frac{\partial S}{\partial H}\right)_T dH \tag{9.8}$$

The heat capacity at constant H is defined by

$$(dQ)_H \equiv C_H\,dT \tag{9.9}$$

Evidently

$$C_H = T\left(\frac{\partial S}{\partial T}\right)_H \tag{9.10}$$

which also follows from (8.22) if P is replaced by its magnetic analog H. Returning to (9.8) we have, upon using (9.7) and (9.10),

$$dQ = C_H\,dT + T\left(\frac{\partial M}{\partial T}\right)_H dH \tag{9.11}$$

which is the magnetic counterpart of (8.23). For an adiabatic process this yields the relation between changes in H and T

$$dT = -\frac{T}{C_H}\left(\frac{\partial M}{\partial T}\right)_H dH \tag{9.12}$$

For paramagnetic materials $(\partial M/\partial T)_H$ is negative. To see this, note that if H is held constant, a temperature increase results in a decrease of M. That is, a temperature increase means greater agitation at the atomic level. Thermal motions tend to destroy the alignment of atomic moments which produce a net magnetic moment M. Thus, higher T means lower M so that $(\partial M/\partial T)_H$ is negative. It follows from (9.12) that dT and dH have the same sign; an adiabatic demagnetization decreases the temperature.

The unusually large magnetic susceptibilities of certain paramagnetic salts make possible the production of ultra-low temperatures. They also provide a means of measuring them. The susceptibility χ is a temperature-dependent quantity and may therefore serve as a thermometer. The paramagnetic salt functions as the core of an inductance coil. The temperature decrease resulting from demagnetization causes an increase in the permeability of the core. This in turn gives rise to an induced emf which is directly proportional to the change in χ.

At sufficiently high temperatures (above 1 K) the susceptibility follows Curie's law

$$\chi = \frac{C}{T} \tag{9.13}$$

where C is the Curie constant for the material, not its specific heat. This simple T dependence of χ does not persist at the lowest operating temperatures. Nevertheless it has become general practice to assume that χ has the Curie form, but with the absolute temperature T replaced by a magnetic temperature T^*. The relation between T^* and T must then be determined experimentally for each paramagnetic material. A technique for converting T^* measurements into T values was invented by Simon and Kurti. Their method, which has always impressed the author as a diabolically clever application of thermodynamics, is discussed later in this section.

....

9.3.2 Show that

$$\left(\frac{\partial M}{\partial T}\right)_H = -\frac{M}{T}$$

for a substance which obeys Curie's law.

....

The low temperature produced by demagnetization is in fact the temperature of the gas of ion magnetic moments, not that of the host lattice. In order for the lattice to be cooled to nearly the same temperature as the ions, two conditions must be satisfied. First, the heat capacity of the lattice must be small

compared to that of the ion moments. This enables the lattice and ion moments to reach equilibrium at a temperature close to the temperature of the moments immediately after the demagnetization. Second, it is necessary that the ion–lattice equilibrium be established quickly, before external heat leaks (such as radiation and vibration) restore equilibrium between the sample and the helium bath. Happily, both of these conditions are satisfied by a number of paramagnetic salts. The lattice specific heat C_{lat} is given by the Debye formula

$$C_{\text{lat}} = 464\left(\frac{T}{\theta}\right)^3 \quad \frac{\text{cal}}{\text{mole-K}} \tag{4.19}$$

where θ is the Debye temperature, a characteristic of the lattice. If we take $\theta \sim 300\,\text{K}$, we find $C_{\text{lat}} \sim 10^{-5}$ cal/mole-K at $T \sim 1\,\text{K}$. The specific heat of the ion moments displays an anomalously large value at some low temperature. The temperature at which the peak occurs is a characteristic of the material, but typical values fall in the 0.01–0.1 K range. The heat capacity of the ion moments may reach 1 cal/mole-K, some 100 million times that of the lattice at the same temperature. This means that even a dilute concentration of ions can cool the lattice without undergoing a serious temperature increase. The high heat capacity of the ion moments also serves as a thermal anchor which maintains the low temperature against heat leaks and, more importantly, can serve as a refrigerant to cool other materials.

The practice of introducing a magnetic temperature T^* merits further discussion. We define T^* in terms of the susceptibility and the Curie constant

$$\chi = \frac{C}{T^*} \tag{9.14}$$

Evidently $T^* = T$ so long as the material obeys Curie's law. The absolute thermodynamic temperature is defined by

$$T = \left(\frac{\partial U}{\partial S}\right) \tag{8.3'}$$

Experimentally, the partial derivative appears as $\Delta U/\Delta S$, the ratio of corresponding changes in U and S. Thus

$$T = \frac{\Delta U}{\Delta S} \tag{9.15}$$

Suppose that the demagnetization measurements indicate a magnetic temperature change of ΔT^*. Let us note that (9.15) may be rewritten as

$$T = \frac{\Delta U/\Delta T^*}{\Delta S/\Delta T^*} \tag{9.16}$$

The temperature T refers to the final ($H = 0$) state. Thus if a means can be devised for evaluating

$$\left(\frac{\Delta U}{\Delta T^*}\right)_{H=0} \quad \text{and} \quad \left(\frac{\Delta S}{\Delta T^*}\right)_{H=0}$$

the absolute temperature can be determined. The numerator of (9.16) is simply the zero-field heat capacity

$$C^* \equiv \left(\frac{\Delta U}{\Delta T^*}\right)_{H=0} \tag{9.17}$$

where C^* is measured by irradiating the sample with a γ-ray source—a neat way of adding energy to a substance surrounded by liquid helium. The change in internal energy equals the energy deposited by the absorbed γ rays

$$\Delta U = Q_\gamma \tag{9.18}$$

The quantity ΔT^* is measured by the change in susceptibility. The rate at which energy is deposited in the sample by the γ rays depends only on the strength of the source, the geometry of the sample, and so on. Hence a determination of the rate at which γ rays deposit energy can be conducted at a higher temperature where the heat capacity is known from calorimetric measurements.

To see how the ratio $\Delta S/\Delta T^*$ is determined, consider Figure 9.9. Plotted there on a T^*–H diagram are the paths of two sequential processes. The

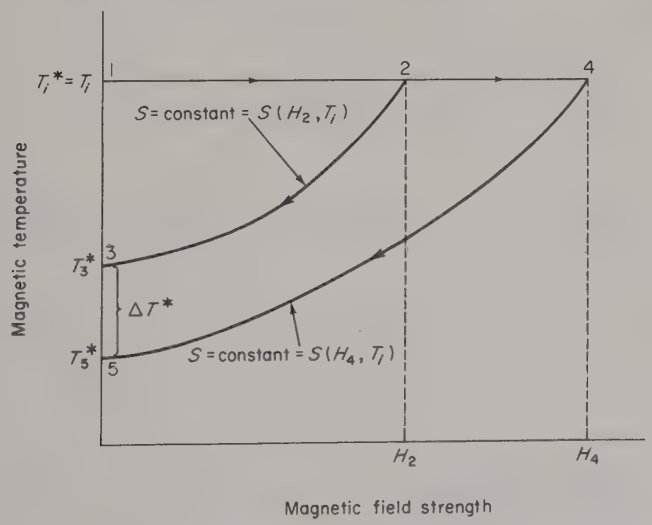


Figure 9.9. T^*–H diagram indicating how the ratio $\Delta S/\Delta T^*$ can be inferred from adiabatic demagnetization data.

processes are identical with those involved in the magnetic cooling sequence depicted by Figure 9.8. Each of the two paths consists of an isothermal magnetization followed by an adiabatic (and isentropic) demagnetization. The temperature at which the isothermal magnetizations proceed is sufficiently high that $T_i^* = T_i$ can be measured by conventional low-temperature thermometry (for example, vapor pressure). The final T^* temperatures, T_5^* and T_3^*, are measured via the susceptibility, so that

$$\Delta T^* = T_3^* - T_5^*$$

is thereby determined. The entropy difference follows from the result (see Problem 9.3.1)

$$S(H, T_i) - S(0, T_i) = -\frac{1}{T_i}\int_0^H H\, dM \tag{9.19}$$

Since T_i is high enough that $T_i^* = T_i$, the susceptibility is given by $\chi = C/T_i$, from which it follows that $M = CVH/T_i$. This leads to the result

$$S(H_4, T_i) - S(0, T_i) = -\frac{CVH_4{}^2}{2T_i{}^2} \tag{9.20}$$

The entropy change required to determine T is

$$\Delta S = S(H_2, T_i) - S(H_4, T_i) = \frac{CV(H_4{}^2 - H_2{}^2)}{2T_i{}^2} \tag{9.21}$$

To recap, the absolute temperature is defined by (8.3) which, experimentally, is equivalent to (9.16). Using (9.17) gives

$$T = \frac{\Delta U/\Delta T^*}{\Delta S/\Delta T^*} = \frac{C^* \Delta T^*}{\Delta S} \tag{9.22}$$

The magnetic temperature follows from measurements of the susceptibility. The heat capacity C^* is determined by using a calibrated γ-ray source and ΔS is inferred from (9.21).

REFERENCES

The liquid–solid transition in nitrogen provides a spectacular demonstration. See:

R. L. Wild and D. C. McCollum, Dramatic Demonstration of Change of Phase, *Am. J. Phys.* **35**, 540 (1967).

A modification of the demonstration is suggested by:

H. C. Jensen, Freezing Nitrogen: A Modification, *Am. J. Phys.* **36**, 919 (1968).

Virtually every aspect of low-temperature physics is discussed at length in:

G. K. White, "Experimental Techniques in Low-Temperature Physics." Oxford Univ. Press, London and New York (1959).

See also:

J. S. Dugdale, "Entropy and Low Temperature Physics," Chapter 15. Hutchinson, London (1966).

A very readable account of cryogenic applications is given in:

R. J. Allen, "Cyrogenics." Lippincott, Philadelphia, Pennsylvania (1964).

Included are discussions of liquefiers, superfluidity, and superconductivity.

An entertaining account of cryogenic principles and applications is given in:

D. K. C. MacDonald, "Near Zero." Doubleday, Garden City, New York (1961).

Historical commentary on magnetic cooling can be found in:

C. G. B. Garrett, "Magnetic Cooling." Harvard Univ. Press, Cambridge, Massachusetts (1954).

Chapter 6 provides much detail on the structures of the paramagnetic salts. See also

C. T. Lane, "Superfluid Physics," Chapter 1. McGraw-Hill, New York (1962).

G. K. White, "Experimental Techniques in Low-Temperature Physics," Chapter 8. Oxford Univ. Press, London and New York (1959).

H. B. G. Casimir, "Magnetism and Very Low Temperatures." Dover, New York (1961).

A detailed development which describes magnetic cooling with superconductors and antiferromagnetic salts as well as with the usual paramagnetic salts is given in:

H. Weinstock, Thermodynamics and Statistical Aspects of Magnetic Cooling, *Am. J. Phys.* **36**, 36 (1968).

A note of caution concerning the reversibility of adiabatic demagnetization is voiced in:

R. Baierlein and B. Bertman, Reversibility in Magnetic Cooling, *Am. J. Phys.* **37**, 101 (1969).

Chapter 10

Elementary Statistical Concepts

Both kinetic theory and statistical mechanics require an appreciation and working knowledge of certain aspects of probability theory. In order not to break the continuity of later analyses we present here some rudimentary notions needed in the development of kinetic theory and statistical mechanics. Additional mathematical machinery required expressly for statistical mechanics will be introduced in Chapter 14. Our immediate needs are minimal compared with what are generally considered to be the elements of probability theory.

10.1 MEAN VALUES AND PROBABILITIES

The physical variables dealt with in practice fall into two classes, discrete and continuous. A discrete variable Z may assume only certain definite values, such as $Z_1, Z_2, Z_3, \ldots$. For example, if Z denotes the number on the face of a die, the possible values are

$$Z_1 = 1, \quad Z_2 = 2, \quad Z_3 = 3, \quad Z_4 = 4, \quad Z_5 = 5, \quad \text{and} \quad Z_6 = 6$$

A continuous variable may take on any value within some range. For example, the speeds of nitrogen molecules in our atmosphere range from zero to thousands of kilometers per second. In this and the immediately following

sections the basic ideas are developed in terms of discrete variables. The corresponding treatment for continuous variables is detailed in Section 10.3.

A quantity of prime importance, both physically and mathematically, is the mean value of a statistical variable. Suppose a series of N trials, that is, measurements of Z, results in N_1 occurrences of Z_1, N_2 occurrences of Z_2, and so on. The mean value of Z for the series of trials is denoted as $\langle Z \rangle$ and is defined by

$$\langle Z \rangle \equiv \frac{Z_1 N_1 + Z_2 N_2 + Z_3 N_3 + \cdots}{N_1 + N_2 + N_3 + \cdots} \tag{10.1}$$

Using summation notation

$$\langle Z \rangle = \frac{\sum_i Z_i N_i}{\sum_i N_i} \tag{10.2}$$

The denominator is just the total number of trials N

$$\sum_i N_i = N$$

Thus,

$$\langle Z \rangle = \frac{\sum_i Z_i N_i}{N} \tag{10.3}$$

....

10.1.1 A series of 28 rolls of a die yields the results $N_1 = 5$, $N_2 = 5$, $N_3 = 5$, $N_4 = 6$, $N_5 = 3$, and $N_6 = 4$, where N_k is the number of times the integer k came up. Show that the mean for the series is approximately 3.32.

....

The Nielsen television ratings and the Gallup political polls sample preferences of the American population. A series of measurements may be thought of as a sample of an indefinitely large assembly—the population of all measurements. In (10.3), N_i/N is the fraction of the sample measurements which give Z_i. The ratio N_i/N is called the relative frequency of Z_i. Experience suggests that the relative frequency approaches a definite limit as the sample size increases. The student may very well have tested this convergence in an error-analysis experiment. The limit of N_i/N as N tends to infinity is called the probability of (the occurrence of) Z_i and is denoted as P_i

$$P_i \equiv \lim_{N \to \infty} \frac{N_i}{N} \tag{10.4}$$

Note that the probabilities are subject to a normalization condition,

$$\sum P_i = 1 \tag{10.5}$$

We instinctively feel that the limit in (10.4) exists, that is, that each N_i increases in such a fashion that N_i/N approaches a definite value. In statistical physics we are more likely to be concerned with predicting the outcome of a contemplated experiment than in describing the results of a completed one. For example, suppose we are asked to predict the mean value of a series of N measurements of a discrete variable Z. We must first assign (predict) the values of N_i/N and then compute $\langle Z \rangle$ using

$$\langle Z \rangle = \frac{\sum_i Z_i N_i}{N} \tag{10.3}$$

Our confidence in the existence of the limit of N_i/N defined by (10.4) influences our prediction. In the absence of any information to the contrary the most plausible guess we can make is that the ratio N_i/N will equal its limiting value P_i. By most plausible we mean that no more plausible estimate of N_i/N could be assigned on the basis of the information provided. The expected (that is, predicted) value of $\langle Z \rangle$ is then

$$\langle Z \rangle = \sum_i Z_i P_i \tag{10.6}$$

It is in this sense that statistical physics utilizes probability theory. The task in statistical physics is to first *predict* the P_i and then to *compute* related mean values using (10.6).

In practice, measurements of a statistical variable Z are used to determine the mean value of some function $Q(Z)$ of the variable. For example, we may wish to determine the mean area of circular disks from measurements of their diameters. If Z denotes the disk diameter,

$$\text{area} \equiv Q(Z) = \frac{\pi Z^2}{4}$$

is the desired physical quantity whose mean value $\langle Q \rangle$ is sought. If a series of N measurements result in N_1 occurrences of Z_1, N_2 occurrences of Z_2, and so on, there are corresponding numbers of occurrences of

$$Q_1 = \frac{\pi {Z_1}^2}{4}, \qquad Q_2 = \frac{\pi {Z_2}^2}{4}, \qquad \cdots$$

The mean value of $Q(Z)$ is thus

$$\langle Q \rangle = \frac{Q_1 N_1 + Q_2 N_2 + \cdots}{N_1 + N_2 + \cdots}$$

or

$$\langle Q \rangle = \sum_i Q_i \frac{N_i}{N} \tag{10.7}$$

If we are asked to predict $\langle Q \rangle$, we will (somehow!) assign probability estimates for the N_i/N, obtaining

$$\langle Q \rangle = \sum_i Q_i P_i \tag{10.8}$$

Evidently, it is immaterial whether we talk of the probability of the occurrence of Z_i or of the probability of the occurrence of $Q_i = Q(Z_i)$. It also should be clear that (10.7) and (10.8) are valid for any physical quantity $Q(Z)$.

Example 10.1 RMS Values

The rms value is defined as the square *root* of the *mean square* value,

$$Z_{\text{rms}} \equiv \sqrt{\langle Z^2 \rangle}$$

To illustrate, we determine the rms value for the die roll data of Problem 10.1.1. The mean square value of Z for the 28 tosses is given by

$$\langle Z^2 \rangle = \frac{\sum_i Z_i^2 N_i}{N}$$

The result is $\langle Z^2 \rangle = 385/28$, which yields $Z_{\text{rms}} = 3.71$. Note that $Z_{\text{rms}} > \langle Z \rangle$, since the larger values of Z carry greater weight in determining $\langle Z^2 \rangle$ in consequence of the squaring operation.

It may happen that the mean value $\langle Z \rangle$ is zero. While this is a nontrivial attribute of the quantity it gives no hint as to the magnitudes of the values assumed by the quantity. The rms value then becomes a useful measure of the magnitudes of the quantity. It would be equally legitimate to deal with the mean magnitude $\langle |Z| \rangle = \sum_i |Z_i| (N_i/N)$. This is seldom done because certain mathematical properties make Z_{rms} much more manageable than $\langle |Z| \rangle$. ✧

....

10.1.2 Prove in general that $Z_{\text{rms}} \geqslant \langle Z \rangle$. [*Hint:* The mean square value of a variable is never negative. Your proof should make use of the fact that $\langle \zeta^2 \rangle \geqslant 0$, where $\zeta \equiv Z - \langle Z \rangle$.]

10.1.3 The standard deviation σ is defined by

$$\sigma \equiv \sqrt{\sum_i (Z_i - \langle Z \rangle)^2 P_i}$$

Show that σ may be expressed as

$$\sigma = \sqrt{\langle Z^2 \rangle - \langle Z \rangle^2}$$

....

Statistical physics enlarges and extends the layman's conception of probability. Everyday usages of the word "probability" are typified by the following situations: With reference to the outcome of one flip of a coin we have the statement, "The probability of heads is $\frac{1}{2}$." With reference to drawing one card from a standard deck one hears statements such as, "The probability of drawing the ace of spades is $\frac{1}{52}$." In connection with the roll of a die we have statements such as, "The probability of throwing a four is $\frac{1}{6}$." Used in these contexts, the probabilities are referred to as *a priori* probabilities. The terminology *a priori* probability refers to the probability *in the absence of prior knowledge*. Thus, for example, suppose you were given a die whose six faces bear the numbers 1–6. The die might be loaded so that the face bearing the "3" turns up in a suspiciously large percentage of the rolls. However, in the absence of any data pertaining to the die's roll characteristics, one would be led to postulate equal probabilities of $\frac{1}{6}$ for the likelihood that a 1, 2, 3, 4, 5, or 6 turn up. Whenever we lack evidence which favors certain possible outcomes over others, it is plausible to assume equal probabilities for all possible outcomes. The acquisition of information may change the probabilities. For example, suppose a ruled line divides a sheet of paper into two parts. You are then told that a single pencil mark dots the paper and asked to assign probabilities for its appearance in either portion. In the absence of any further information the most plausible assignment of probabilities is $\frac{1}{2}$ for either side. If you receive the additional information that the area on one side of the ruled line is twice that on the other side, the most plausible assignment of probabilities becomes $\frac{1}{3}$ and $\frac{2}{3}$.

•••••• ••••••

10.1.4 An urn contains six white balls and six black balls.

(a) What is the probability that a ball drawn from the urn will be white?
(b) If a white ball is drawn from the urn (and not returned), what is the probability that a second draw will yield another white ball?

10.1.5 An underwater cable 300 km long breaks.

(a) What is the probability that the break occurred within 2 km of the middle?
(b) What is the probability that the break occurred within 2 km of a specified end?
(c) What is the probability that the break occurred within 2 km of a point located 1 km from a specified end?

10.1.6 The serial number of a one dollar bill has eight digits. A complete run of bills consists of serial numbers 00000001–99999999.

(a) What is the probability that all eight digits are 9's?
(b) What is the probability that the last two digits are 9's?
(c) What is the probability that all eight digits are the same (11111111, 22222222, ...)? Note that there is no bill with eight zeros.

•••••• ••••••

The assignment of equal *a priori* probabilities in situations like the die tossing example illustrates Laplace's Principle of Insufficient Reason:

> **In the absence of evidence to the contrary, all possible outcomes are equally probable.**

It was the great French mathematician Laplace who stated that "probability theory is nothing but common sense reduced to calculation." When we assign *a priori* probabilities we are using common sense. Thus if you are asked for the probability of drawing an ace from a standard deck of cards, your common sense answers 4/52. On the other hand, if you have evidence which favors one outcome over another, you certainly would not assign them equal probabilities. For example, if you know that one deuce and one king have been removed from the deck, common sense and very little calculation leads to a probability of 4/50 for drawing an ace, but only 3/50 for the chance of drawing a deuce.

Example 10.2 Star–Star Collisions

To illustrate the power of common sense probability arguments, we bring them to bear on a problem related to the collision of two galaxies. Galaxies come in two shapes. Young ones tend to have a pinwheel (or spiral arm) structure. Old ones are usually ellipsoidal. For our purposes we invent a third type, cubical in shape. Such liberties as this may introduce uncertainties of a factor of two or three into the final result. This is within the spirit of the calculation; we are after only a rough estimate.

We suppose that two cubical galaxies meet head-on and pass through each other. The question we ask is this: How many collisions between stars can be expected? To obtain a definite answer we need some astronomical data. Astronomy reveals that most galaxies are about the same size as our own. The characteristic dimension of a galaxy is 10^4 pc, where 1 pc (parsec) is close to 3×10^{18} cm.† Furthermore there appear to be roughly the same number of stars, about 10^{11}, in most galaxies. For our estimate we shall take each galaxy to be a cube of side 10^4 pc, containing 10^{11} stars.

† One parsec is approximately equal to 3 light years.

The average radius of a star is also needed. Our own star, the Sun, is nearly typical, its radius being 7×10^{10} cm. To simplify the arithmetic we assume that each star has a radius of 3×10^{11} cm $= 10^{-7}$ pc. We further assume that the stars are spread randomly throughout their cubical volumes.

Because of the powerful attraction of gravity two stars can disrupt each other without actually touching. We demand a first class smash—we count only collisions which would occur in the absence of gravitational attraction. This requires that the distance between the centers of the two stars be less than the diameter. Equivalently we can double the radius of all the stars in one galaxy and regard all those in the other as points or, perhaps more picturesquely, as darts. One galaxy becomes a set of 10^{11} identical targets. The other becomes a swarm of darts. In addition to the collective motions which carry the galaxies together there are stellar motions within the galaxies. These motions render each star in one galaxy a potential target for each of the dart stars.

We first ascertain the chance that one of the 10^{11} darts strikes a star during its trip through the target galaxy. This is the same as asking for the probability that a dart strike one of 10^{11} targets spread over an area of 10^8 pc^2, that is, we can place all of the target stars on one face of the cube since the incident star has the opportunity to travel across the entire galaxy. This (common sense!) probability is simply the fraction of the area of the face covered by target stars. The target area of each star is $\pi(2 \times 10^{-7})^2$ pc^2, or about 10^{-13} pc^2. With 10^{11} target stars the total target area is 10^{-2} pc^2, so that

$$\begin{aligned}\text{probability that any one incident star will strike a target star} &= \frac{10^{-2}\ \text{pc}^2}{10^8\ \text{pc}^2} \\ &= 10^{-10}\end{aligned}$$

However, there are 10^{11} incident stars, each having the same chance of colliding with a target star. The expected number of star–star collisions therefore works out to 10. Thus even with 10^{11} stars shooting at another 10^{11}, only about 10 collisions are likely—there is a lot of empty galaxy out there. ✧

The predictive aspects of statistical physics are typified by the following problem: A crystal contains N atoms. There are many different possible energies open to each atom. We label the possible energies $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_M$. The number of possible energies M is enormous. Let N_i denote the number of atoms with energy ε_i. Determine the most probable values of the N_i subject to the constraints.

$$\sum_{i=1}^{M} N_i = N \qquad \text{(total number of atoms is } N\text{)}$$

and

$$\sum_{i=1}^{M} \varepsilon_i N_i = E \qquad \text{(total energy of atoms is } E\text{)}$$

The first constraint means only that the number of atoms must be the same for all proposed schemes of distribution. The second condition might refer to the physical situation in which the crystal is isolated. In such a case the total energy of the atoms E is constant. Not every conceivable way of assigning energies to the N atoms is consistent with $\sum \varepsilon_i N_i = E$. From an algebraic viewpoint there are two equations and M unknowns, and $M \gg 2$. Clearly the N_i are not uniquely determined by the two equations of constraint. Since the number of equations is far fewer than the number of unknowns we require some principle —a guiding light—which lets us assign values to the N_i. Our later study of statistical mechanics will reveal the guiding light to be none other than the second law of thermodynamics. In short, the scheme by which statistical mechanics would have the atoms distributed is such that the entropy of the crystal is maximized.

It is important to understand the significance of the N_i which statistical mechanics "cranks out." These N_i are not the measured numbers of atoms having energy ε_i. Such a direct measurement is not possible. The value of N_i assigned by statistical mechanics is in effect a prediction of the number of atoms which possess energy ε_i. This prediction is made in compliance with all constraints ($\sum N_i = N$; $\sum \varepsilon_i N_i = E$; entropy maximized). There is no claim that a crystal satisfying the constraints will in fact have the prescribed distribution of atoms. What statistical mechanics says is that if

$$\sum N_i = N \qquad \text{and} \qquad \sum \varepsilon_i N_i = E$$

and if the entropy is a maximum, then the most likely distribution of the atoms is such and such.†

This situation is comparable to saying that the *a priori* probability of rolling a four is $\frac{1}{6}$. There is no guarantee that a four will turn up in one-sixth of the rolls, but it is most likely that this will occur—no other outcome is more likely.

Suppose that we are given the maximum entropy prediction of the N_i. Then $P_i = N_i/N$ is the probability that a crystal atom have an energy ε_i. With the P_i in hand we can proceed to compute mean values according to

$$\langle Q \rangle = \sum_{i=1}^{M} Q_i P_i, \qquad Q_i = Q(\varepsilon_i) \qquad (10.8')$$

† The actual form of N_i is not important here. Later in the game we will discover that N_i is proportional to $\exp(-\varepsilon_i/kT)$, where T is the absolute temperature and k is the Boltzmann constant.

This $\langle Q \rangle$ is the expected (predicted) mean value of the quantity Q. In practice one computes the mean values of observable quantities so that the predicted value of $\langle Q \rangle$ can be compared with experiment.

10.2 CONDITIONAL PROBABILITIES

In many instances it is necessary to determine the probability for compound events or propositions. For example, the *a priori* probability for heads on the flip of a coin is $\frac{1}{2}$. Suppose we wish to determine the probability that heads will turn up on two consecutive flips. The set of two flips comprise a compound event. To cite another example, suppose it is known that two aces have been removed from a standard deck of cards. (Which two is unknown.) The problem: What is the probability of drawing the ace of spades? The proposition before us is compounded of two elementary propositions: (a) the ace of spades is in the deck and (b) the ace of spades will be drawn, given that it is in the deck. In the discussion which follows we use the following notation: Capital letters are used to designate propositions (events). Thus, H might denote the proposition, "The flip of a coin results in heads." The corresponding probability is designated as $P(\mathrm{H})$:

$$P(\mathrm{H}) = \text{probability of heads}$$

More generally,

$$P(\mathrm{A}) = \text{probability that event A occurs}$$

The product AB denotes the compound proposition that both A and B occur and $P(\mathrm{AB})$ is the probability that both A and B occur. A basic problem is this: How is $P(\mathrm{AB})$ related to $P(\mathrm{A})$ and $P(\mathrm{B})$? To establish such a relation requires the introduction of conditional probabilities. As the name implies, the conditional probability for some event refers to a probability based on prior knowledge—a condition which affects its likelihood. The notation $P_{\mathrm{B}}(\mathrm{A})$ denotes "the probability of A, given that B has occurred."

As an example of a conditional probability, consider the following situation. Let A denote the proposition "an ace will be drawn" and let B denote the proposition "two aces are replaced by two deuces," both statements in reference to a standard deck of 52 cards. The *a priori* probability that an ace will be drawn is 4/52,

$$P(\mathrm{A}) = \frac{4}{52}$$

However, if it is known that two deuces have replaced a pair of aces, the

probability of selecting an ace is reduced to 2/52. Thus the conditional probability $P_B(A)$ is

$$P_B(A) = \frac{2}{52}$$

....

10.2.1 If two aces are replaced by two deuces (B) and if C denotes the proposition "a deuce will be drawn," what is $P_B(C)$?

....

We can now state the relation between $P(AB)$, $P_B(A)$, and $P(A)$, namely

$$P(AB) = P(B) \cdot P_B(A) \tag{10.9}$$

This relation is in fact a definition of the conditional probability $P_B(A)$. To illustrate the use of (10.9), consider a deck of cards from which two aces have been removed, leaving 50 cards. Which two aces remain is not known. Let A denote the proposition, "the ace of spades will be drawn" and let B denote the proposition "the ace of spades will be in a deck from which two aces have been removed." Since we do not know which two aces remain we have $P(B) = 2/4$ for the probability that the ace of spades is in the partial deck of 50 cards. If we know that the ace of spades is present, the probability of drawing it is 1/50

$$P_B(A) = \frac{1}{50}$$

Thus using (10.9)

$$P(AB) = P(B)P_B(A) = \frac{1}{100}$$

....

10.2.2 Determine the conditional probability that a six will turn up on the roll of a die, given that an even number turns up.

10.2.3 An urn contains six white balls and six black balls. What is the probability that two draws will yield two white balls when:

(a) the first ball drawn is returned to the urn before the second draw;
(b) the first ball drawn is not returned to the urn before the second draw.

....

A very special and important use of (10.9) arises where the occurrence of A does not depend at all on the occurrence of B. If the probability for A is not

affected by B, we have

$$P_B(A) = P(A)$$

whereupon (10.9) becomes

$$P(AB) = P(A) \cdot P(B) \tag{10.10}$$

If the probability for A is not affected by B, we say that A and B are independent. Thus for independent events, A and B, the joint probability $P(AB)$ is simply the product of $P(A)$ and $P(B)$ as expressed by (10.10). As a very simple example consider a pair of flips of a coin. Let

$$A = \text{heads on first flip}$$

and

$$B = \text{heads on second flip}$$

Then

$$AB = \text{heads on both flips}$$

Since $P(A) = P(B) = \frac{1}{2}$, (10.10) gives the common sense result $P(AB) = \frac{1}{4}$.

We will have occasion to apply (10.10) in Section 10.4. Conditional probabilities are required in many areas of statistical physics.

10.3 CONTINUOUS VARIABLES AND DISTRIBUTION FUNCTIONS

In the case of continuous statistical variables it is unwise (from an operational viewpoint) to speak of the probability of a value of Z. Any series of trials will record a finite number of values of Z, yet there exist a nondenumerable (uncountable) infinity of possible values. Hence the probability (number of occurrences divided by number of trials) is zero for all but a few values of Z—those that turned up.

For example, consider the task of determining the probability that a 10-year-old boy in Denmark has a specified height. We can imagine a horde of doctors and nurses invading schools and homes throughout Denmark, measuring and recording the heights of all 10-year-old boys. For each observed height there could be assigned a probability, namely,

$$P_i = \frac{\text{number of 10-year-old Danish boys of height } h_i}{\text{total number of 10-year-old Danish boys}}$$

This would assign nonzero values of P_i only to those heights observed. This approach would make P_i zero for all heights not observed. Certainly, the fact that the heights 1.384 m, 1.386 m, and 1.388 m were observed, but not the

heights 1.385 m, 1.387 m, and 1.389 m, does not suggest the impossibility of the latter heights. Suppose that instead of P_i we were to specify the fraction of 10-year-old Danish boys whose heights fall in a specified range of heights. The result would be a description much more in line with our intuitive notions about the likelihood of observing any particular height.

To this end we find it both meaningful and useful to introduce a probability *distribution function* for the values of a statistical variable. The distribution function measures the likelihood that a value of a statistical variable falls within a specified range of values. Let $dP(Z)$ denote the probability that the value lies in the range from Z to $Z+dZ$. We define the probability distribution function $F(Z)$ by

$$dP = F(Z)\,dZ \tag{10.11}$$

In terms of the distribution function we can give $dP(Z)$ a very helpful geometrical interpretation. Figure 10.1 displays the form typical of several distribution functions.

From Figure 10.1 we see that $dP(Z) \simeq F(Z)\,dZ$ equals the area of the strip of width dZ between the abscissa and the graph of $F(Z)$. This interpretation is extremely useful. For example, we can immediately conclude that the probability that Z lie in the range from Z_A to Z_B is given by the total area under the $F(Z)$ plot from Z_A to Z_B. Formally,

$$\text{probability that } Z \text{ lies in the range from } Z_A \text{ to } Z_B = \int_{Z_A}^{Z_B} F(Z)\,dZ \tag{10.12}$$

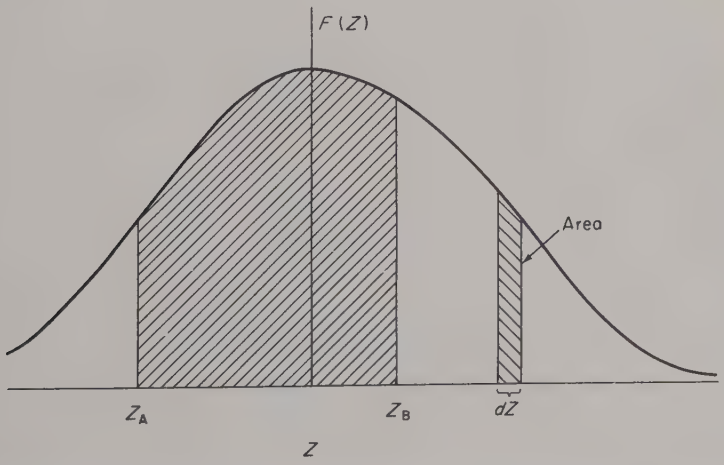


Figure 10.1. Probability is measured by the area beneath a plot of the distribution function (area $= F(Z)\,dZ = dP$). The area between Z_A and Z_B gives the probability that a measurement of Z will fall in that range.

In the discrete variable case, the sum of all probabilities totaled unity

$$\sum_i P_i = 1 \tag{10.5}$$

For continuous variables the summation is replaced by an integration over the full range of possible values of Z. Thus the analog of (10.5) for continuous variables is the normalization condition

$$\int_{Z_L}^{Z_U} dP(Z) = \int_{Z_L}^{Z_U} F(Z)\, dZ = 1 \tag{10.13}$$

where Z_L and Z_U designate the lower and upper limits which Z can attain. Geometrically, (10.13) is simply a statement of the fact that the total area beneath the plot of $F(Z)$ equals unity.

In general one can make the transition from discrete to continuous probability formulas by replacing a summation over P_i by an integration over dP. In particular, the mean value equation

$$\langle Q \rangle = \sum_i Q(Z_i)\, P_i \tag{10.8'}$$

converts to

$$\langle Q \rangle = \int_{Z_L}^{Z_U} Q(Z)\, dP \tag{10.14}$$

....

10.3.1 The distribution function for a continuous variable x is

$$F(x) = A(1-x^2), \qquad -1 \leqslant x \leqslant 1$$

$$F(x) = 0, \qquad |x| > 1$$

(a) Show that $A = \frac{3}{4}$ is required for a properly normalized distribution function $\int_{-\infty}^{+\infty} F(x)\, dx = 1$.

(b) Evaluate $\langle x \rangle$ and $\langle x^2 \rangle$.

(c) What can be said about $\langle x^N \rangle$, where N is an odd integer? Mental calculation is in order rather than the pencil-and-paper variety.

....

An especially important distribution function is encountered in the kinetic theory of gases. There is a distribution of molecular speeds in a gas. The corresponding probability distribution function is denoted $g(v)$. We write

$$dP(v) = g(v)\, dv \tag{10.15}$$

where $dP(v)$ is the probability that the speed of a molecule lies in the range from v to $v+dv$. We could equivalently state that $g(v)\, dv$ is the fraction of the molecules with speeds in the range v to $v+dv$.

The Maxwellian† form of the molecular distribution function is

$$g_M(v) = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp\left(\frac{-mv^2}{2kT}\right) \tag{10.16}$$

The Maxwellian form of $g(v)$ describes the distribution of speeds in a gas in thermal equilibrium at a temperature T. The quantity k is the Boltzmann constant and m is the mass of a molecule. The numerical factor in $g_M(v)$ guarantees that the probability $dP(v)$ is normalized to unity, that is, that

$$\int_{\text{all speeds}} dP(v) = \int_{v=0}^{\infty} g_M(v)\, dv = 1 \tag{10.17}$$

....

10.3.2 Verify that $\int_{v=0}^{\infty} g_M(v)\, dv = 1$. Refer to Appendix A for assistance in evaluating the definite integral.

10.3.3 (a) Show that the mean molecular speed for the Maxwellian distribution is

$$\langle v \rangle = \sqrt{\frac{8kT}{\pi m}}$$

(b) Show that the rms speed for the Maxwellian distribution is

$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}}$$

The relevant integrals are furnished in Appendix A.

(c) By writing $k = R/N_0$ and noting that the molar mass is $M = N_0 m$ show that these results may be expressed as

$$\langle v \rangle = \sqrt{\frac{8RT}{\pi M}}, \qquad v_{\text{rms}} = \sqrt{\frac{3RT}{M}}$$

10.3.4 A tabletop is ruled with parallel lines spaced 1 in. apart. Toothpicks, each 1 in. in length, are dropped onto the tabletop in random fashion.

(a) What is the probability $T(\theta)$ that a toothpick crosses (or touches) a line, given that it makes an angle θ with the lines?

(b) The angle θ is a random variable over the range $\theta = 0$ to $\theta = \pi/2$, that is, the method of dropping does not favor any particular angle. What is

† Named after James Clerk Maxwell, who first deduced its form through theoretical arguments.

the probability $dP(\theta)$ that a toothpick makes an angle with the lines which lies in the range from θ to $\theta+d\theta$? (It is also possible to consider the variable θ as ranging from 0 to π, or even from 0 to 2π, but be careful of signs if you make such a choice. Remember, probabilities cannot be negative.)

(c) Show that the mean probability $\langle T \rangle$ that a toothpick cross or touch a line equals $2/\pi$. This is the fraction of toothpicks which cross or touch a line. [*Note:* Assuming that a large number of trials gives the assumed random distribution, this gives a statistical method for determining π. Thus if N drops result in M crossings or touchings, we have $M/N = 2/\pi$ or $\pi = 2N/M$. Try it yourself with, say, 157 toothpicks. A favorable outcome would yield 100 crossings or touchings.]

....

10.4 FREE PATH LENGTHS AND THEIR DISTRIBUTION

The atoms of a gas are incessantly in motion. Collisions between atoms cause deflections and result in irregular zig-zag paths. The distance between collisions is called the *free path length*. Collisions occur randomly so that the length of individual free paths cannot be predicted. A useful measure of free path length is the average or mean distance between collisions, the so-called *mean free path* λ. We can estimate λ as follows: Let the effective radius of an atom be denoted by R_0. By effective we mean that a collision occurs if the centers of two atoms approach to within $2R_0$ or less. It is customary to denote the effective cross-sectional area of an atom as σ. If we take $R_0 \simeq 2\times 10^{-8}$cm, the typical cross section is

$$\sigma \equiv \pi {R_0}^2 \simeq 10^{-15}\ \text{cm}^2$$

As the atoms speed about between collisions they sweep out cylindrical volumes. Over a free path length r, an atom sweeps out a volume σr. The mean volume swept out is $\sigma\lambda$, where λ is the mean free path. Each atom is eagerly vying for its share of the container volume. The mean volume $\sigma\lambda$ is therefore one atom's share of the container volume. With N atoms in a volume V we have

$$\sigma\lambda = \text{volume per atom} = \frac{V}{N} \tag{10.18}$$

Under standard conditions, 1 mole of a gas occupies a volume of 22.4 liters and contains Avogadro's number of atoms. The volume per atom is therefore

$$\frac{V}{N} = \frac{22{,}400\ \text{cm}^3}{6.023\times 10^{23}} \simeq 4\times 10^{-20}\ \text{cm}^3$$

It follows that $\lambda \simeq 4 \times 10^{-5}$ cm. Note that λ is about 2000 times as great as the effective atomic diameter. Thus under ordinary conditions the mean free path is gigantic by comparison with atomic dimensions. A gas atom spends the majority of its time in free flight.

The distance between successive collisions cannot be predicted. However, a probability for a given free path length can be assigned. We write

$$W(r) = \begin{array}{c}\text{probability that particle will travel a distance } r \\ \text{without suffering a collision}\end{array} \tag{10.19}$$

The distance of travel is measured from the last collision. The collisions occur at random. At random means that there is no rule or law which can specify how far any given atom will travel before suffering a collision. Translated into mathematical terms, at random means that the probability that a collision occur between r and $r+dr$ is proportional to dr, but independent of r.

$$\begin{array}{c}\text{probability that a collision will occur} \\ \text{between } r \text{ and } r+dr\end{array} = \alpha\, dr \tag{10.20}$$

The proportionality constant α generally depends on factors such as gas pressure, temperature, and composition. The important point is that α does not depend on the distance r.

In traveling a distance dr the particle either does or does not suffer a collision. Thus

$$\begin{array}{c}\text{probability that a collision does not} \\ \text{occur between } r \text{ and } r+dr\end{array} = 1 - \alpha\, dr \tag{10.21}$$

It is now possible to write two equations for $W(r+dr)$, the probability that the particle travel a distance $r+dr$ without undergoing a collision. For sufficiently small dr, $W(r+dr)$ can differ only slightly from $W(r)$. Denoting this change by dW we write,

$$W(r+dr) = W(r) + dW \tag{10.22}$$

A second equation for $W(r+dr)$ follows by noting that a collision-free path of length $r+dr$ consists of two independent steps; one of length r followed by one of length dr. As independent events, the overall probability $W(r+dr)$ is just the product of the separate probabilities.

$$W(r+dr) = \begin{bmatrix}\text{probability that a particle} \\ \text{will travel a distance } r \text{ without} \\ \text{suffering a collision}\end{bmatrix} \cdot \begin{bmatrix}\text{probability that no} \\ \text{collision will occur} \\ \text{between } r \text{ and } r+dr\end{bmatrix} \tag{10.23}$$

The first factor is $W(r)$, while the second is given by (10.21). Thus

$$W(r+dr) = W(r)[1-\alpha\, dr] \tag{10.24}$$

Equating (10.22) and (10.24) leads to

$$\frac{dW}{W} = -\alpha\, dr \tag{10.25}$$

which may be integrated to give

$$W(r) = Ae^{-\alpha r} \tag{10.26}$$

....

10.4.1 The constant A equals one, being fixed by the condition $W(r=0)=1$. Explain the physical significance of this condition.

....

Setting $A = 1$ gives the final result,

$$W(r) = e^{-\alpha r} \tag{10.27}$$

To check (10.27) and pave the way for determining the physical significance of α we introduce

$$dP(r) \equiv \text{probability of a free path length in the range } r \text{ to } r+dr \tag{10.28}$$

A free path length between r and $r+dr$ also consists of two independent events, a collision-free trip of length r followed by a collision between r and $r+dr$. Thus

$$dP = \begin{bmatrix}\text{probability that a particle}\\ \text{will travel } r \text{ without}\\ \text{suffering a collision}\end{bmatrix} \cdot \begin{bmatrix}\text{probability that a}\\ \text{collision will occur between}\\ r \text{ and } r+dr\end{bmatrix} = W\alpha\, dr$$

Using (10.27),

$$dP(r) = e^{-\alpha r}\alpha\, dr \tag{10.29}$$

Knowledge of $dP(r)$ enables one to determine the physical significance of α. The mean free path, designated earlier by λ, follows from (10.14) with $Q = r$

$$\lambda \equiv \langle r \rangle = \int_{r=0}^{\infty} r\, dP(r) = \int_{r=0}^{\infty} re^{-\alpha r}\alpha\, dr \tag{10.30}$$

An integration by parts shows that the integral equals $1/\alpha$. Thus

$$\text{mean free path} = \lambda = \frac{1}{\alpha}$$

The relevant probabilities, $W(r)$ and $dP(r)$, may now be rewritten in terms of λ.

$$W(r) = e^{-r/\lambda}$$

$$dP(r) = \frac{e^{-r/\lambda}\, dr}{\lambda} \tag{10.31}$$

The exponential factor $e^{-r/\lambda}$ shows that there is but a slim chance of a free path length many times λ, the mean value.

....

10.4.2 (a) An atom in a gas suffers collisions at random. Show that the probability $W(t)$ that the atom travel for a time t (measured from time of last collision) without undergoing a collision is given by

$$W(t) = e^{-\nu t}$$

where ν is some constant.

(b) Show that $\nu = 1/\tau$, where τ is the mean free time, the mean time between collisions. The definition of τ is

$$\tau \equiv \langle t \rangle = \int_0^\infty t \, dP(t)$$

where $dP(t)$ is the probability that the atom travel without a collision for a time t and then suffer a collision between t and $t+dt$.

....

REFERENCES

There are many fine textbooks in probability theory. The following contain introductory chapters especially suitable for the beginner:

B. V. Gnedenko and A. Ya. Khinchin, "An Elementary Introduction to the Theory of Probability." Dover, New York (1961).

J. V. Upensky, "Introduction to Mathematical Probability." McGraw-Hill, New York (1937).

V. E. Gmurman, "Fundamentals of Probability Theory and Mathematical Statistics." American Elsevier, New York (1968).

A recent review of Maxwell's kinetic theory research is contained in:

S. G. Brush, James Clerk Maxwell and the Kinetic Theory of Gases: A Review Based on Recent Historical Studies. *Am. J. Phys.* **39**, 631 (1971).

Chapter 11

Gases : Models and Equipartition

In this chapter we introduce a model of a gas. The model pictures a gas as a vast collection of particles—atoms and molecules—which move about incessantly. This model leads to the ideal gas equation of state and provides a mechanical interpretation of temperature.

The degree of freedom concept is developed. Crudely stated, the number of degrees of freedom equals the number of independent ways in which the system may store energy.

The equipartition of energy theorem is presented; the theorem states that the energy of a system is distributed equally among the degrees of freedom, that is, there is an equal partitioning of energy.

We apply the equipartition of energy theorem to the ideal gas model. This leads to a prediction of the heat capacities, C_p and C_v, which prove to be in good agreement with the experimental values.

Finally, we consider deviations from the ideal behavior. The concept that pressure measures energy per unit volume is introduced and used to deduce the van der Waals and virial equations of state.

11.1 THE IDEAL GAS

We have previously noted that all gases may be described by the same equation of state, provided that the pressure is not too high and the temperature

not too low. This is the by now familiar, ideal gas equation of state. For n moles of gas, the pressure, temperature, and volume are related by

$$PV = nRT \tag{11.1}$$

where R is the universal gas constant whose value depends on the system of units adopted

$$\begin{aligned} R &= 8.317 \times 10^7 \text{ erg/K} \\ &= 8.317 \text{ joule/K} \\ &= 1.987 \text{ cal/K} \\ &= 0.08208 \text{ liter-atm/K} \end{aligned}$$

An alternate form of the equation of state is also useful, namely,

$$PV = NkT \tag{11.2}$$

in which N is the total number of molecules present in the volume V, and $k = R/N_0$ is the Boltzmann constant.

Example 11.1 Atomic Size versus Container Size

We cannot see the molecules which comprise the air we breathe. Atoms and molecules presumably are small. How small? What is the size of a typical atom? To answer this question does not require an electron microscope or an understanding of quantum mechanics. A reliable estimate of atomic size can be based on the observation that solids and liquids are virtually incompressible. Extremely high pressures are required to alter the volume of a liquid by a few percent. Many solids can withstand very large compressive stresses without significant distortion. By contrast, a gas is readily compressed. These facts suggest that the atoms in a gas are relatively far apart while those in liquid or solid form are packed rather tightly. As we now demonstrate, the dimension characteristic of atoms is 10^{-8} cm. This minute length is called an angstrom and abbreviated as 1 Å.

To estimate atomic sizes we assume that the volume of a solid or liquid equals the total volume of the constituent atoms. That is, we choose to ignore the space between atoms. This seems a reasonable assumption, for if there were much space between atoms it should be possible to squeeze them closer together by applying pressure. However, as already noted, squeezing solids and liquids scarcely affects their volumes.

Consider 1 mole of a substance. Let V denote its volume and let M signify the gram-molecular weight. The number of molecules in 1 mole is Avogadro's number $N_0 \simeq 6 \times 10^{23}$. Next, we envision the volume V as composed of N_0 subvolumes of equal size—one for each molecule. More precisely, we picture

each molecule as residing in a cubical room. If the length of the edge of such a cube is d, its volume is d^3. With such a picture d becomes a measure of atomic size. Taken collectively the N_0 cubes comprise the total volume V. Thus

$$V = N_0 d^3$$

Next we observe that the mass density ρ may be expressed by

$$\rho = \frac{\text{mass of 1 mole}}{\text{volume of 1 mole}} = \frac{M}{V} = \frac{M}{N_0 d^3}$$

Solving for d gives

$$d = \left(\frac{M}{\rho N_0}\right)^{1/3} \tag{11.3}$$

Provided M is in grams and ρ is in grams per cubic centimeter we find

$$d \simeq \left(\frac{5M}{3\rho}\right)^{1/3} \text{Å}, \qquad 1 \text{ Å} = 10^{-8} \text{ cm} \tag{11.3a}$$

As Table 11.1 shows, the factor $(5M/3\rho)^{1/3}$ is essentially the same for many elemental solids and liquids. Other materials furnish similar results. In view of these results we are led to conclude that the angstrom characterizes atomic

TABLE 11.1

Characteristic Atomic Dimensions

	Element	M (gm)	ρ (gm/cm^3)	$\left(\frac{5M}{3\rho}\right)^{1/3}$ (Å)
Solids	Li	6.9	0.53	2.8
	Ne	20.2	1.00	3.2
	Na	23.0	0.97	3.4
	Al	27.0	2.70	2.6
	K	29.1	0.86	4.2
	Ar	39.9	1.65	3.4
	Zn	65.4	7.1	2.5
	Rb	85.5	1.53	4.5
	Ag	107.9	10.5	2.6
	Cs	132.9	1.87	4.9
	Pt	195.2	21.4	2.5
	Au	197.0	19.3	2.6
	Pb	207.2	11.3	3.1
Liquids	He	4.0	0.12	3.8
	Ne	20.2	1.21	2.9
	Ar	39.9	1.39	3.6
	Kr	83.8	2.61	3.7
	Xe	131.3	3.06	4.1
	Hg	200.6	13.6	2.9

sizes. The corresponding atomic volume is roughly 10^{-24} cm^3. Let us see how this atomic volume compares with the container volume.

One mole of atoms in the gaseous state under standard conditions occupies a volume of 22.4 liters, about 2×10^4 cm^3. The number of atoms in the molar volume is $N_0 \simeq 6 \times 10^{23}$. If each atom has a volume of 10^{-24} cm^3, the total atomic volume of 1 mole of atoms is close to 1 cm^3. Comparing this with the container volume of over 2×10^4 cm^3 we conclude that atoms in the gaseous state are dispersed throughout a volume more than 10,000 times as great as their atomic volume. Stated another way, if all of the atoms were piled in one corner of their container, they would fill only about 1/10,000 of its volume. ✧

The next step is to put forth a model of a gas. Consideration of the equation of state for the model leads to an important interpretation: The temperature of a gas is a measure of the average kinetic energy of its atoms and molecules.

11.2 IDEAL GAS MODEL AND CRITIQUE

The following set of four assumptions define a particular model of a gas. We shall show that this model leads to the ideal gas equation of state. A critique of the model follows, indicating the experimental basis for each assumption and the extent to which each is realistic. Our intention of deducing the equation of state implicitly assumes that the gas is in thermal equilibrium with the container walls.

The Model:

(a) The gas consists of a large number of particles (atoms†) which obey Newton's laws of motion.

(b) The motion of the atoms is random.

(c) The volume of the atoms themselves is negligible by comparison with the volume of the container. The atoms of a monatomic gas are treated as point masses.

(d) The atoms collide elastically with the container walls, but do not collide with one another. For a monatomic gas, this means that the atoms possess only translational kinetic energy, except during the instant of impact in a wall collision.

Critique of the Model:

(a) Experimental evidence for assuming that the gas contains many atoms is furnished by Perrin's determinations of Avogadro's number. In a variety

† For the present we consider a monatomic gas.

of experiments, Perrin established that $N_0 \simeq 6 \times 10^{23}$. At a moderate pressure, even a fraction of a cubic centimeter of gas contains an enormous number of atoms.

Choosing to operate at the level of classical mechanics is motivated by an important historical fact. Kinetic theory successfully evolved within a classical framework. Newton's laws of motion are adequate for the development of most aspects of the theory. Of course, atoms "really" obey the laws of quantum mechanics. Indeed, certain paradoxes arise in kinetic theory if one insists on applying classical mechanics at every turn. In particular, the heat capacities of many gases exhibit certain anomalies which may be explained only through a quantum mechanical treatment. The heat capacity anomalies are historically significant for they sounded early warnings that something was wrong with the foundations of 19th century physics.

For the moment we concentrate on the successes of kinetic theory as they stemmed from the framework of classical physics. Later we will find out what went wrong. Still later we will see how quantum physics can correct shortcomings of the classical theory.

(b) The assumption of random motion is quite general and would be incorporated in more sophisticated models, whether classical or quantum mechanical. The experimental evidence for random motion is based on observations of Brownian motion. The earliest observations of Brownian motion involved colloidal suspensions and demonstrated the random motion of atoms in the liquid state. The random motion of atoms in a gas was demonstrated by Fletcher and Millikan in 1911. They observed the Brownian motion of electrically charged oil droplets, using many of the same techniques developed by Millikan to determine the charge of the electron.

(c) Only in a few gross respects does an atom behave like a point mass. However, it is these same few characteristics which are most crucial for a determination of the equation of state. The translational kinetic energy of the atom can be written as $\frac{1}{2}mv^2$, where m is the atomic mass and v is the speed of the center of mass. The linear momentum is given by mv. With respect to translational kinetic energy and linear momentum, any structure of the atom is irrelevant—it behaves like a point mass. However, as pointed out earlier, atoms do take up space, an effect which becomes more important as the gas is compressed. For rarified gases, the atomic volume effect can be accounted for by a little "handwaving." One simply replaces the container volume V by $V-V_{ex}$, where V_{ex} is the so-called excluded volume. The idea is that the entire container volume is not available to any given atom; the presence of the other atoms excludes a certain volume. This point is developed further in Section 11.6.

Liquid-to-vapor transition data justify the neglect of the atomic volume in a rarified gas. For example, 1 gm of water at the normal boiling point occupies

a volume of approximately 1 cm^3. After vaporization it occupies a volume of nearly 1700 cm^3, suggesting that each molecule in the vapor has about 1700 times as much room to move about in as it had in the liquid state. In this respect, the datum for water is typical.

(d) The assumption of elastic collisions is artificial. A realistic picture is much more complicated. Some atoms collide with the wall and adhere to it (adsorption). However, the assumption of elastic collisions does not lead to any error so long as the gas is in thermal equilibrium with the container walls. Thermal equilibrium requires that no net transfer of energy occur between the gas and the walls. The elastic collision assumption rules out *any* energy transfer between the gas and the walls, automatically ensuring that no net energy is transferred.

Perhaps the most significant assumption is that the atoms do not collide with one another. Even in a rarified gas, atoms experience mutual interatomic forces. These interatomic forces depend largely on the structure of the atom. Thus, for example, the force between a pair of hydrogen atoms is distinctly different from the force between a pair of helium atoms. As later analysis will show, it is these interatomic forces which give rise to the nonideal characteristics of real gases. We deal with this shortcoming of our model in Section 11.6.

11.3 EQUATION OF STATE

Having chosen a model, the next step is to establish the equation of state. This requires that we compute the gas pressure, the average normal force per unit area arising from the collisions of the atoms with the container walls. The gas is enclosed in a rectangular container having sides of length L_x, L_y, and L_z.

Let v_x denote the x component of velocity of an atom which subsequently collides with the wall (see Figure 11.1). The atom rebounds with an x

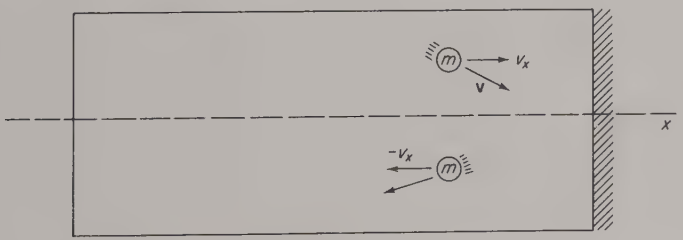


Figure 11.1. An atom undergoes an elastic collision with the wall. The normal component of its velocity changes sign. Its speed is unchanged.

component of velocity equal to $-v_x$. The change in its x component of momentum is

$$(p_x)_{\text{aft}} - (p_x)_{\text{bef}} = -2mv_x$$

with aft standing for after and bef for before. A change of equal magnitude and opposite direction is experienced by the container wall. The x momentum transferred to the wall is

$$\begin{matrix}\text{momentum transfer to wall as a} \\ \text{result of one collision}\end{matrix} \equiv \Delta p_x = 2mv_x \tag{11.4}$$

This momentum is parceled out among the many atoms comprising the container walls. The momentum transferred by a single atom is not observable—measuring instruments respond to the combined effect of many such collisions.

An atom with velocity component v_x travels the length L_x of the container in a time L_x/v_x. The time Δt between collisions with a particular wall is twice as great

$$\Delta t = \frac{2L_x}{v_x} \tag{11.5}$$

The average rate at which such an atom transfers momentum to the wall is

$$\frac{\Delta p_x}{\Delta t} = \frac{2mv_x}{2L_x/v_x} = \frac{mv_x^2}{L_x} \tag{11.6}$$

By Newton's second law, the rate of change of momentum equals force

$$\frac{\Delta p_x}{\Delta t} = f_x \tag{11.7}$$

Thus the average force exerted by the impacts of an atom having an x-velocity component v_x is

$$f_x = \frac{mv_x^2}{L_x} \tag{11.8}$$

The total force on the wall F_x is obtained by summing over all particles.

$$F_x = \sum_{\text{a.a.}} f_x = \sum_{\text{a.a.}} \frac{mv_x^2}{L_x} \tag{11.9}$$

where a.a. stands for all atoms. The area of the container wall normal to the x axis is $A = L_y L_z$. The pressure P on this area is

$$P = \frac{F_x}{A} = \sum_{\text{a.a.}} \frac{mv_x^2}{L_x L_y L_z} \tag{11.10}$$

Since $L_x L_y L_z$ is the volume V of the container we can write

$$PV = \sum_{\text{a.a.}} mv_x^2 \tag{11.11}$$

Now, the pressure of the gas is the same on every wall; there is no preferred direction of motion. Thus we also have the results

$$PV = \sum_{\text{a.a.}} mv_y^2 = \sum_{\text{a.a.}} mv_z^2 \tag{11.12}$$

Further, since

$$v_x^2 + v_y^2 + v_z^2 = v^2 \tag{11.13}$$

we can write

$$PV = \tfrac{1}{3} \sum_{\text{a.a.}} m(v_x^2 + v_y^2 + v_z^2) = \tfrac{1}{3} \sum_{\text{a.a.}} mv^2 \tag{11.14}$$

The total kinetic energy K_{tot} of the gas is

$$K_{\text{tot}} = \sum_{\text{a.a.}} \tfrac{1}{2}mv^2 \tag{11.15}$$

whereupon we have

$$PV = \tfrac{2}{3}K_{\text{tot}} \tag{11.16}$$

To maintain the particle point of view we write K_{tot} as N times the mean kinetic energy per atom $\langle K \rangle$

$$K_{\text{tot}} = N\langle K \rangle$$

and thence,

$$PV = \tfrac{2}{3}N\langle K \rangle \tag{11.17}$$

As it stands, (11.17) is the equation of state which the model has produced.

At this juncture it is worthwhile to note that the analysis applies equally well to diatomic and polyatomic gases. Only considerations involving the translational kinetic energy and linear momentum of the center of mass enter the pressure calculation. Since we follow only the center of mass motion, any structure of the atoms or molecules is irrelevant in so far as the equation of state is concerned. Thus all ideal gases obey the same equation of state regardless of any microscopic structure attributed to their atoms and molecules. This conclusion is in accord with the experimental data for real gases. Experimentally, the equation of state obeyed (approximately) by all gases is

$$PV = NkT \tag{11.2}$$

Comparing (11.17) and (11.2) we are led to conclude that the temperature of an ideal gas is proportional to the mean kinetic energy of the atoms.

Specifically, we must have

$$\langle K \rangle = \langle \tfrac{1}{2}mv^2 \rangle = \tfrac{3}{2}kT \tag{11.18}$$

if the model is to reproduce the empirical form of the equation of state. Note that (11.18) is *not a definition of temperature*, since we cannot measure the mean kinetic energy of an atom, that is, such a definition would not be operational. Our model of the ideal gas does present us with a microscopic interpretation of temperature. It takes us beyond the semiquantitative statement "equality of temperature is a condition for thermal equilibrium," and gives us a mechanical feel for temperature. The temperature of a gas measures the translational kinetic energy of the atomic motions. Clearly this revelation would never be forthcoming from any macroscopic thermodynamic analysis. In passing we note that the random atomic motions are referred to the container. Thus, for example, the temperature of the gas in a balloon is not altered by transporting it in an auto at 60 mph, in a jet at 600 mph, or in a spacecraft at 6000 mph.

....

11.3.1 (a) Compute the mean kinetic energy of a monatomic gas atom for a temperature of 300 K. Express the result in electron volts. (1 eV = 1.6×10^{-12} erg.) Note that the mean energy depends on the temperature of the gas, but not on the mass of the atoms.

(b) Compute the rms speed of a helium atom for a temperature of 300 K. It will be recalled that the rms value of a quantity is the square root of the mean value of the square of the quantity. Thus

$$v_{\text{rms}} = \sqrt{\langle v^2 \rangle}$$

The mass of a helium atom is 6.66×10^{-24} gm.

11.3.2 One of the attractive features of "nearby" outer space is the essentially complete vacuum existing there. A total vacuum is not achieved because of the solar wind. Vast quantities of ionized hydrogen—electrons and protons—are spewed off the Sun. This material streams outward, filling the solar system with a stream of particles. The number density and rms speed of the solar wind protons have been measured by Mariner 2 and other space probes. The values are

$$n_{\text{p}} \simeq 3/\text{cm}^3, \qquad v_{\text{rms}} \simeq 500 \text{ km/sec}$$

These data refer to regions "near" Earth, but at altitudes well above its permanent atmosphere.

(a) Compute the corresponding kinetic pressure of the solar wind, assuming that the protons behave as an ideal gas [see (11.17)].

(b) How does the solar wind pressure compare with the best man-made vacuum achieved on Earth (approximately 10^{-10} Torr)?

....

At this point, kinetic theory lets us reap another benefit from the model. In our thermodynamic treatment, the ideal gas was defined by two relations. The first was the equation of state (11.1). The second relation specified the internal energy to be (for n moles),

$$U = nC_v T \tag{11.19}$$

with the molar heat capacity C_v being a constant. In Section 11.5 we show that our model leads to this same relation. Furthermore, the model predicts the value of C_v. Here then, kinetic theory will make its first modest prediction, thereby allowing a comparison with experiment. It is first necessary to introduce the notion of degrees of freedom and to formulate the equipartition of energy theorem.

11.4 THE EQUIPARTITION OF ENERGY

Several examples can best serve to illustrate the idea of a *degree of freedom*. The system depicted in Figure 11.2 consists of a bead (regarded as a point mass)

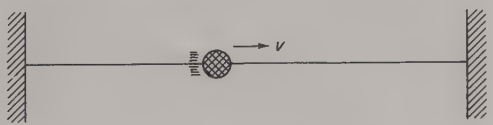


Figure 11.2. Bead on a wire—a system with one degree of freedom.

of mass m which is free to move along a horizontal wire. The constraints imposed by the wire permit only one-dimensional motion. The total energy of the bead is kinetic,† $E_{\text{bead}} = \frac{1}{2}mv^2$. Only one variable, the speed of the bead, is needed to specify the energy of the bead. Note that E is proportional to v^2; we say that E is a quadratic function of v. We also say that the bead possesses one degree of freedom—its total energy depends on but one variable. In general the

† We take the gravitational potential energy of the bead to be zero. Other choices could be made. The essential point is that the gravitational potential is a constant, independent of the position or speed of the bead.

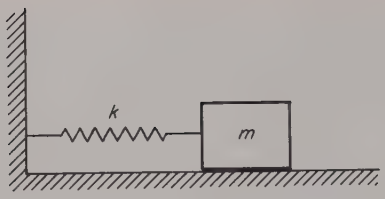


Figure 11.3. The simple harmonic oscillator has two degrees of freedom.

number of degrees of freedom possessed by a system† equals the number of independent variables needed to specify its energy.

An often-used system in physics is the simple harmonic oscillator (SHO). Figure 11.3 shows one realization of the SHO. The spring–mass system is capable of undergoing one-dimensional motion—oscillations along the x axis. The total energy of the one-dimensional SHO is

$$E = \tfrac{1}{2}mv^2 + \tfrac{1}{2}kx^2$$

As such the SHO possesses two degrees of freedom. Two independent variables, v and x, are needed to specify the kinetic and potential energies of the oscillator. We shall use f to denote the number of degrees of freedom of a system, so that $f = 2$ for the SHO. The SHO is an extremely important system throughout physics. At the elementary level it is useful in illustrating Newtonian dynamics and the essentials of oscillatory motion. In quantum mechanics it turns out that certain systems (for example, electromagnetic waves and sound waves) behave essentially like collections of SHO's.

Figure 11.4 shows a system with six degrees of freedom. It is a three-dimensional SHO. The mass is free to move in any direction from its equilibrium position. Let x, y, and z denote the components of its displacement and v_x, v_y, and v_z the corresponding velocity components. The total energy of the system is

$$E = \tfrac{1}{2}mv_x^2 + \tfrac{1}{2}mv_y^2 + \tfrac{1}{2}mv_z^2 + \tfrac{1}{2}k_x x^2 + \tfrac{1}{2}k_y y^2 + \tfrac{1}{2}k_z z^2$$

so that $f = 6$ for the three-dimensional SHO. Note that the total energy is a quadratic function of each of the six independent variables x, y, z, v_x, v_y, and v_z.

The atoms of a crystalline solid can be visualized as the masses of a three-dimensional SHO. The springs represent interatomic forces which bind the atoms together and give the solid its elastic quality. A macroscopic chunk of such a solid thus appears as a gigantic bedspring-like arrangement with masses (atoms) at the intersections of the spring network.

Another important system is the free point mass—a point mass subject to no

† Here the word "system" refers to the component parts (for example, atoms, molecules, and so on) which comprise a thermodynamic system.

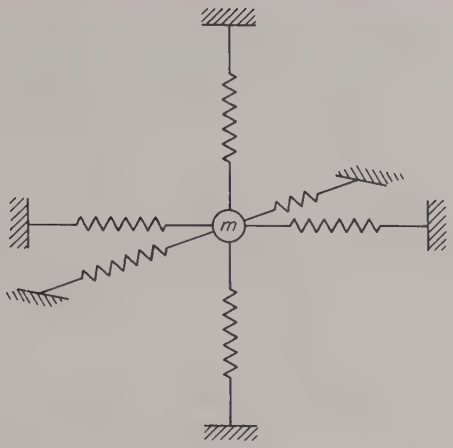


Figure 11.4. A three-dimensional oscillator possesses six degrees of freedom.

external forces. The free point mass has already served us as a model for the monatomic gas atom. The total energy of the point mass is

$$E = \tfrac{1}{2}mv_x^2 + \tfrac{1}{2}mv_y^2 + \tfrac{1}{2}mv_z^2$$

Three independent variables, v_x, v_y, and v_z, are required to specify the total energy so that $f = 3$ for the free point mass.

If the point mass experiences a uniform gravitational acceleration g in the negative z direction, it acquires an additional degree of freedom. The total energy is

$$E = \tfrac{1}{2}mv_x^2 + \tfrac{1}{2}mv_y^2 + \tfrac{1}{2}mv_z^2 + mgz$$

with mgz being the gravitational potential energy. Unlike the previous examples E is not a quadratic function of all four independent variables.

Another system of interest is the so-called dumbbell molecule. This model pictures a diatomic molecule as a pair of point masses bound together by a rigid rod. As Figure 11.5 suggests, the dumbbell molecule appears to be a

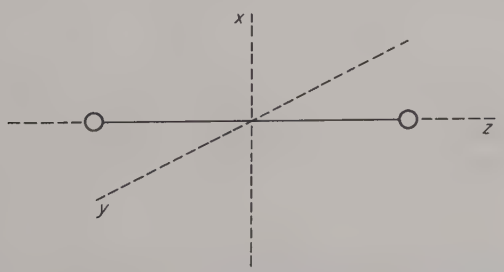


Figure 11.5. The dumbbell molecule has five degrees of freedom.

miniaturization of the circus strongman's equipment. For the dumbbell molecule we have $f = 5$. Three degrees of freedom are accounted for by the translational kinetic energy of the molecule. If m is the total mass of the molecule and v_x, v_y, and v_z are the center of mass velocity components, the translational kinetic energy is

$$E_{\text{trans}} = \tfrac{1}{2}mv_x^2 + \tfrac{1}{2}mv_y^2 + \tfrac{1}{2}mv_z^2$$

The dumbbell molecule is free to rotate about its center of mass and thus can possess rotational energy as well. Taking the z axis to lie along the rod we can write

$$E_{\text{rot}} = \tfrac{1}{2}I_x\omega_x^2 + \tfrac{1}{2}I_y\omega_y^2$$

where I_x and I_y are the moments of inertia about the x and y axes while ω_x and ω_y are the corresponding angular velocities about these axes. The assumption of point masses means that $I_z = 0$. The total energy is the sum of the translational and rotational contributions

$$E_{\text{dumb}} = \tfrac{1}{2}mv_x^2 + \tfrac{1}{2}mv_y^2 + \tfrac{1}{2}mv_z^2 + \tfrac{1}{2}I_x\omega_x^2 + \tfrac{1}{2}I_y\omega_y^2$$

The variables v_x, v_y, v_z, ω_x, and ω_y are independent so that $f = 5$ for the dumbbell molecule.

As a final example we consider a somewhat more realistic model of a diatomic molecule. There is a vast amount of spectroscopic evidence to support the reality of the rotational degrees of freedom ascribed to the dumbbell molecule. Additionally, there is a wealth of spectroscopic data suggesting that a diatomic molecule also possesses vibrational degrees of freedom. The two atoms in the molecule are not bound rigidly, but can move relative to one another in a vibratory fashion. These vibrational capabilities can be incorporated by using a model in which the rigid rod of the dumbbell is replaced by a stiff spring (see Figure 11.6). The resulting system—a vibrating dumbbell molecule—has seven degrees of freedom. In so far as the vibrational motions are concerned, the molecule behaves like a one-dimensional SHO, which has

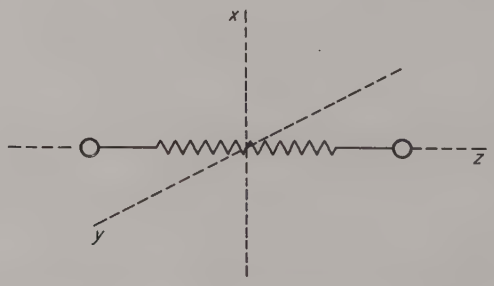


Figure 11.6. The vibrating dumbbell molecule has two vibrational degrees of freedom.

two degrees of freedom. Therefore, the vibrational capability adds two degrees of freedom to the three translational and two rotational degrees of freedom of the rigid dumbbell model. Furthermore, the vibrational energy E_{vib} can be expressed in the form

$$E_{\text{vib}} = \tfrac{1}{2}m'v_\zeta^2 + \tfrac{1}{2}k\zeta^2$$

where m' is a mass (a particular combination of the masses of the two atoms) and k is an effective spring constant describing the electronic force which binds the atoms. The independent variables ζ and v_ζ can be related to the displacements and velocities of the two atoms. We need not indicate these relations since the only fact germane to our later analysis is that E_{vib} is a quadratic function of both ζ and v_ζ.

These examples serve to illustrate the notion of a degree of freedom. A system possesses one degree of freedom for each independent variable needed to determine the total energy. Except for the gravitational potential energy encountered in one example, the total energy proved to be a quadratic function of these independent variables. The equipartition theorem is concerned with the relation between the temperature and the mean energy associated with each degree of freedom.

In Section 11.3 we noted the fact that there was no preferred direction of motion for the atoms so that

$$\langle \tfrac{1}{2}mv_x^2 \rangle = \langle \tfrac{1}{2}mv_y^2 \rangle = \langle \tfrac{1}{2}mv_z^2 \rangle \tag{11.20}$$

Since

$$\langle \tfrac{1}{2}mv_x^2 \rangle + \langle \tfrac{1}{2}mv_y^2 \rangle + \langle \tfrac{1}{2}mv_z^2 \rangle = \langle \tfrac{1}{2}mv^2 \rangle = \tfrac{3}{2}kT \tag{11.18'}$$

it follows that equal mean energies are ascribed to each of the three degrees of freedom. Each degree of freedom contributes $\frac{1}{2}kT$ to the total mean energy

$$\langle \tfrac{1}{2}mv_x^2 \rangle = \langle \tfrac{1}{2}mv_y^2 \rangle = \langle \tfrac{1}{2}mv_z^2 \rangle = \tfrac{1}{2}kT \tag{11.21}$$

This applies on a per particle basis. The energy per mole attributable to each degree of freedom is

$$N_0(\tfrac{1}{2}kT) = \tfrac{1}{2}RT$$

The equipartition of energy theorem states that each degree of freedom contributes $\frac{1}{2}kT$ to the mean energy. Thus if the particles (atoms or molecules) of a gas possess f degrees of freedom, the mean energy per particle is $\frac{1}{2}fkT$. The thermal energy of the system is partitioned equally among the degrees of freedom—thus the name of the theorem.

The theorem is subject to a number of restrictions. It is true only for those degrees of freedom which contribute quadratic terms to the total energy. Thus, for example, the point mass in a uniform gravitational field has four degrees of

freedom, but only three contribute quadratic terms to the energy,

$$E = \tfrac{1}{2}mv_x^2 + \tfrac{1}{2}mv_y^2 + \tfrac{1}{2}mv_z^2 + mgz$$

The mean energy is

$$\langle E \rangle = \tfrac{3}{2}kT + \langle mgz \rangle$$

The equipartition theorem decrees that each of the three translational degrees of freedom contributes equally to $\langle E \rangle$, but is silent on the relation between $\langle mgz \rangle$ and the temperature. Statistical mechanics does enable one to relate $\langle mgz \rangle$ and T, but has little to say, in a general sense, about those degrees of freedom which do not contribute quadratic terms to the energy.

For future reference we note that for both monatomic and diatomic ideal gases, each of the degrees of freedom contributes a quadratic term to the total energy so that the mean energy of an ideal gas atom or molecule may be written

$$\langle E \rangle = \tfrac{1}{2}fkT \tag{11.22}$$

....

11.4.1 The equipartition energy is $\frac{1}{2}kT$ per particle.

(a) Prove that the corresponding energy per mole is $\frac{1}{2}RT$, where R is the gas constant.

(b) Assuming that an equipartition of energy prevails, what is the total energy of 1 mole of an ideal gas (temperature T) whose molecules behave like rigid dumbbells.

....

11.5 INTERNAL ENERGY AND HEAT CAPACITIES

The internal energy is taken to be the total energy of the particles (atoms or molecules) comprising the gas. Two sources contribute to the internal energy of a real gas. First, each particle possesses its own private energy. For monatomic gases this energy is wholly kinetic. Polyatomic molecules may also possess rotational and vibrational energy. The second source of internal energy in a real gas is the mutual potential energy of the particles—generally referred to as an interaction energy to stress the fact that two or more particles are involved. The interaction energy is electrical in nature. It arises because the charges of neighboring atoms and molecules exert mutual Coulomb forces. In a rarified gas, these forces and the associated potential energies are ordinarily negligible because of the overall electrical neutrality of the particles. However, when two atoms collide, that is, pass within a few angstroms of each other, their electron atmospheres clash and the Coulomb forces and potential energies become significant. The mean distance which an atom travels between collisions in a

rarified gas may be several thousand angstroms. Comparing this distance with the collision range of a few angstroms shows that the atoms spend most of their time in free flight. Only rarely are they interacting with other atoms. It was this fact which justified ignoring collisions in the ideal gas model. To ignore collisions is completely equivalent to neglecting the interaction energy. When this is done the internal energy of an ideal gas is just the sum of the private energies carried by each particle. We can now formulate these ideas quantitatively and use the equipartition theorem to deduce the internal energy of an ideal gas.

Let $\langle E \rangle$ denote the mean energy of a particle and let E_{INT} denote the interaction energy of the entire gas. The internal energy U is

$$U = N\langle E \rangle + E_{\text{INT}} \tag{11.23}$$

where $N = nN_0$ is the number of particles in the gas. For an ideal gas, E_{INT} is ignored so that we have simply

$$U = N\langle E \rangle, \qquad \text{ideal gas} \tag{11.24}$$

The mean energy is related to the temperature by the equipartition theorem. For particles with f degrees of freedom,

$$\langle E \rangle = \tfrac{1}{2} fkT \tag{11.22}$$

whence,

$$U = \tfrac{1}{2} fNkT \tag{11.25}$$

An alternate form for U is obtained by noting that $Nk = nR$, whence

$$U = n\left(\frac{fR}{2}\right)T \tag{11.26}$$

In our thermodynamic dealings with the ideal gas the form assumed for internal energy was

$$U = nC_v T \tag{11.19}$$

Comparison of (11.19) with (11.26) shows that the *molar* heat capacity predicted by our model is

$$C_v = \frac{fR}{2} \tag{11.27}$$

....

11.5.1 Show that (11.27) follows from the thermodynamic relation (4.24)

$$C_v = \left(\frac{\partial U}{\partial T}\right)$$

....

Remembering that $R = 1.987$ cal/K we arrive at the handy form

$$C_v \simeq f \text{ cal/K} \tag{11.27a}$$

that is, each degree of freedom contributes approximately 1 cal/K to C_v. As we showed earlier C_p and C_v are related (for an ideal gas) by

$$C_p = C_v + R$$

so that the model gives

$$C_p = \left(\frac{f+2}{2}\right)R \tag{11.28}$$

The ratio of the heat capacities

$$\frac{C_p}{C_v} = \gamma \tag{11.29}$$

is also subject to a check. The quantity γ may be determined from measurements of the speed of sound in the gas.† From (11.27) and (11.28) it follows that for an ideal gas γ has the value

$$\gamma = \frac{f+2}{f} \tag{11.30}$$

Note that these results are deductions—predictions—coaxed from the kinetic theory model. For a monatomic gas, $f = 3$ so that the model predicts

$$C_p = \frac{5}{2}R = 4.98 \quad \frac{\text{cal}}{\text{K}} \tag{11.31}$$

$$\gamma = \frac{5}{3} = 1.67$$

For the diatomic dumbbell molecule ($f = 5$) we obtain

$$C_p = \frac{7}{2}R = 6.97 \quad \frac{\text{cal}}{\text{K}} \tag{11.32}$$

$$\gamma = \frac{7}{5} = 1.40$$

---

† The relation between γ and the speed of sound is

$$\text{(sound speed)} = \sqrt{\frac{\gamma P}{\rho}}$$

where P is the pressure and ρ the mass density of the gas.

TABLE 11.2

Molar Heat Capacity C_p and γ for Monatomic and Diatomic Gases[a]

Gas	C_p (cal/K)	γ
Helium (He)	5.00	1.66
Argon (A)	5.00	1.67
Neon (Ne)	—	1.64
Krypton (K)	—	1.68
Xenon (Xe)	—	1.66
Hydrogen (H_2)	6.87	1.41
Carbon monoxide (CO)	6.94	1.40
Nitrogen (N_2)	6.94	1.40
Oxygen (O_2)	7.03	1.40
Hydrogen chloride (HCl)	7.07	1.40
Chlorine (Cl_2)	8.15	1.36

[a] The values refer to a pressure of 1 atm and a temperature near 15 °C. Adapted with permission from E. H. Kennard, "Kinetic Theory of Gases." McGraw-Hill, New York (1938).

The vibrating dumbbell molecule ($f = 7$) yields the results

$$C_p = \frac{9}{2}R = 9.96 \quad \frac{\text{cal}}{\text{K}} \tag{11.33}$$

$$\gamma = \frac{9}{7} = 1.28$$

Table 11.2 summarizes the results of measurements for several gases. The values quoted are for measurements made at a pressure of 1 atm and temperatures near 300 K. Agreement between experiment and theory is excellent for the monatomic gases. The monatomic gas values of C_p and γ exhibit only very slight variation with temperature. The results for diatomic gases tell a sadder tale. The experimental values are close to those predicted for either the dumbbell model or the vibrating dumbbell model. However, the table does not reveal a most disturbing fact; namely that C_p exhibits a definite temperature variation. This is completely at variance with the ideal gas models—all of which predict temperature-independent values of C_p and γ. Figure 11.7 shows the temperature variation of C_p for molecular hydrogen H_2. Hydrogen exhibits a heat capacity characteristic of a monatomic gas at low temperatures. At higher temperatures it shows a value of C_p near $(7/2)R$, the value predicted for an ideal gas of dumbbell molecules. At still higher temperatures it appears that

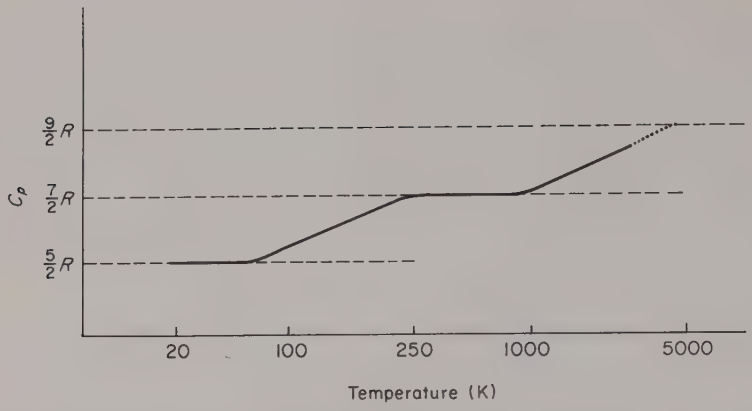


Figure 11.7. The molar heat capacity of molecular hydrogen showing the apparent excitation of additional degrees of freedom as the temperature rises. The temperature scale is logarithmic. The high-temperature data suggest—but do not confirm—that C_p is tending toward (9/2) R. A significant fraction of the molecules dissociate before the temperature reaches 5000 K.

C_p is approaching the value of (9/2) R, appropriate for the vibrating dumbbell model. The hydrogen molecules appear to acquire additional degrees of freedom—even pieces of them—as the temperature increases.

Try as they might, physicists could not reconcile the temperature variation of the heat capacity with any sort of mechanical model. By the beginning of the 20th century there was a gradual awakening. Physicists began to explore the possibility that it was the *physics, not the model*, which was in need of modification. Their efforts culminated in the development of quantum mechanics and quantum statistical mechanics. Thus the rules of the game—the laws of Newtonian mechanics—were found to have their limitations. At the atomic level, new rules were called for—the laws of quantum mechanics. The models have come through unscathed. Indeed, the dumbbell and vibrating dumbbell models are successfully used today in conjunction with quantum statistical mechanics. The quantum statistical formalism explains the observed temperature variation of C_p for H_2 and for other diatomic molecules.

....

11.5.2 Figure 11.8 shows a plot of the molar heat capacities of several crystalline solids versus absolute temperature. Observe that the plots of C_v tend toward a value of roughly 6 cal/mole-K with increasing temperature.

(a) How many degrees of freedom must be ascribed to each atom in a solid if it is to have a molar heat capacity of approximately 6 cal/mole-K?

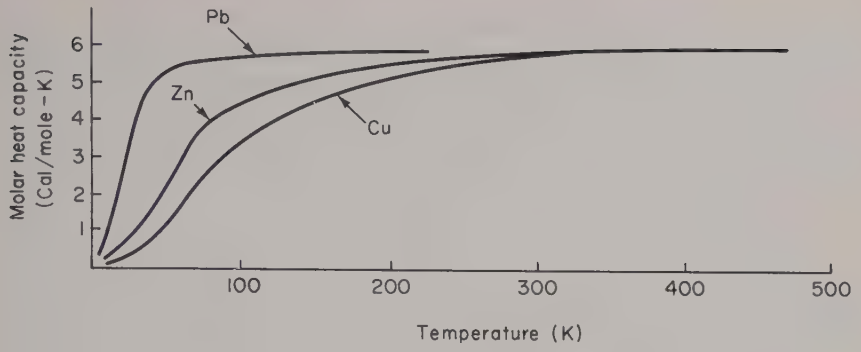


Figure 11.8. Molar heat capacities of three solids.

(b) What model(s) of a solid would accomplish this? Note the rapid decline in C_v at low temperatures. This behavior cannot be understood in terms of the equipartition of energy. However, the quantum theory again comes to the rescue and provides a theory of specific heats which is in excellent agreement with the low-temperature experimental results.

11.5.3 Rigid adiabatic walls surround a system which consists of 1 mole of a gas and 1 mole of a solid. The gas is ideal and monatomic. The atoms of the solid behave like three-dimensional oscillators, that is, they possess six degrees of freedom. Initially a rigid adiabatic wall separates the gas and solid. The initial temperatures are $T_s = 150\,\text{K}$, $T_g = 300\,\text{K}$. The adiabatic wall separating the solid and gas is then removed. Determine the final equilibrium temperature of the system.

11.5.4 We have constructed the equation of state of an ideal gas by consideration of atoms bouncing about inside a container. The resulting ideal gas equation of state was

$$PV = NkT = \tfrac{2}{3}U, \qquad U = \tfrac{3}{2}NkT$$

Electromagnetic radiation (light waves) exhibits particle characteristics. These particles are called *photons*. Photons display the following properties:

(a) A photon travels in a vacuum with the speed of light c.

(b) The photon has an energy ε given by $\varepsilon = h\nu$, where ν is the frequency of the source of radiation and h is Planck's constant.

(c) The photon has a momentum p whose magnitude is $p = \varepsilon/c$. The photon momentum is parallel to its velocity, just as for ordinary particles. Using these properties show that the equation of state for a photon gas is given by

$$PV = \tfrac{1}{3}U$$

where P is the photon (light) pressure, V is the volume of the container, and U is the internal energy of the photon gas. Your calculation should mimic the procedure followed in deriving the ideal gas equation of state. In particular you are to assume that the photons suffer elastic collisions with the walls of the container. [*Hint:* Recall that $A_x B_x + A_y B_y + A_z B_z = \mathbf{A}\cdot\mathbf{B}$ and note that the momentum and velocity of the photon are parallel and satisfy

$$\mathbf{p}\cdot\mathbf{v} = pc = \varepsilon$$

11.5.5 Show that the equation of state of an ideal gas can be expressed in the form

$$P = (\gamma - 1)\frac{U}{V}$$

Note that this form reveals pressure as a measure of internal energy density U/V.

....

11.6 PRESSURE, ENERGY DENSITY, AND DEPARTURES FROM IDEAL GAS BEHAVIOR

The ideal gas model has one serious deficiency—it takes no account of intermolecular forces. As noted earlier, neglecting intermolecular forces is equivalent to ignoring the mutual potential energy of the molecules—the so-called interaction energy. Taking account of the interaction energy leads to a more realistic equation of state. In the process we uncover a new and useful aspect of pressure.

If we compare (11.2) and (11.25), we see that[†]

$$PV = \tfrac{2}{3}U \tag{11.34}$$

In terms of the internal energy density (internal energy per unit volume)

$$u \equiv \frac{U}{V} \tag{11.35}$$

we have

$$P = \tfrac{2}{3}u \tag{11.36}$$

The fact that pressure and energy density are proportional is most significant.

[†] Throughout this discussion we confine ourselves to monatomic gases.

Indeed, if we were to insist on interpreting pressure only as a force per unit area, we might be led to such *erroneous* conclusions as "there is a pressure on the container walls but nowhere else within the gas," or—*equally incorrect*—"a gas not confined by material walls has no pressure." Such erroneous ideas can be squelched by adopting the viewpoint that

pressure measures energy density.

This point of view is extremely fruitful in considerations involving systems which possess energy in different forms. The energy density associated with each form of energy makes a contribution to the total pressure. For example, a large cloud of interstellar hydrogen gas may contract under the action of its own gravitational forces. The system possesses energy of three distinct types. The hydrogen atoms possess kinetic energy by virtue of their motion. The gravitational forces responsible for the contraction also give rise to a gravitational potential energy. Such a protostar also radiates as it contracts, that is, the hydrogen atoms emit radiant energy. The radiant energy constitutes a photon gas and makes a contribution to the total pressure of the system.

To amplify our plea that pressure be regarded as a measure of energy density we show that the pressure acting on a charged conducting surface just equals the energy density of the electrostatic field. The energy density associated with an electric field E is

$$u_E = \frac{E^2}{8\pi} \tag{11.37}$$

Consider a conducting surface carrying a net charge. The charges distribute themselves over the surface in a fashion dependent upon its shape, but in general they try to get away from one another.† In moving away from each other the charges literally stretch the conducting surface. They make it a little bit larger, thereby decreasing the surface charge density and the energy stored in the electric field. We wish to show that the electrostatic force per unit area—the pressure resulting from the surface charge—is precisely $E^2/8\pi$, that is, that pressure measures energy density. To calculate this pressure, or electrical stress as it is sometimes referred to, we pick a small element of the surface. It follows from Gauss' law that the electric field strength just above the surface is

$$E_{\text{abo}} = \frac{4\pi q}{A} \tag{11.38}$$

† More precisely, they spread out so as to minimize the total energy stored in the electric field which they create.

where q is the charge on the element and A is its area. Because we deal with a static field and a conducting surface, the field just below the surface, that is, inside the conductor, is zero.

$$E_{\text{bel}} = 0$$

The field just outside the area A is a superposition of two fields, the field E_q produced by the charge q and the field E' set up by the other charges, not on A. The same is true for the field just below the surface. It follows from the fact that $E_{\text{bel}} = 0$ that

$$E' = E_q = \tfrac{1}{2}E_{\text{abo}}$$

Below the surface E_q has the same magnitude, but opposite direction, and the net field is zero. The field E' exerts a force on the charge q which occupies the area A. This force is

$$F = qE' = \tfrac{1}{2}qE_{\text{abo}}$$

It is this force which stretches the surface. The electrical pressure is

$$P_E = \frac{F}{A} = \frac{1}{2}\frac{qE_{\text{abo}}}{A}$$

However, from (11.38), $q/A = (1/4\pi)E_{\text{abo}}$, whence

$$P_E = \frac{1}{8\pi}E_{\text{abo}}^2 = u_E \tag{11.39}$$

that is, the electrostatic pressure just equals the energy density in the electrostatic field.

A similar result holds for magnetic fields, the energy density in a field B being

$$u_B = \frac{B^2}{8\pi} \tag{11.40}$$

There is a magnetic pressure, equal to u_B, which occasionally manifests itself in spectacular fashion by tearing apart the windings of an electromagnet.

The pressure of a light wave, the so-called radiation pressure, may be related to the combined electric and magnetic field energy densities. It may also be interpreted in terms of photon impacts (see Problem 11.5.4).

As an application of the "pressure measures energy density" concept we turn to the problem of accounting for the nonideal behavior of gases. For purposes of comparison we need the ideal gas equation of state

$$PV = nRT \tag{11.1}$$

which may be rewritten in the form

$$\frac{Pv}{kT} = 1 \tag{11.41}$$

in which $v = V/N$ is the specific volume, the volume per atom.

Two important equations of state, frequently used to represent the behavior of real gases, are the van der Waals equation

$$\left(P + \frac{a}{v^2}\right)(v-b) = kT \tag{11.42}$$

and the virial equation of state

$$\frac{Pv}{kT} = 1 + \frac{B}{v} + \frac{C}{v^2} + \cdots \tag{11.43}$$

The coefficients B, C, ... are termed the second, third, ... virial coefficients. The virial coefficients can be determined experimentally and it is found that they vary with temperature. In a rarified gas the volume per atom v is large enough that successive terms in the virial equation become progressively smaller. The terms B/v, C/v^2, and so on, measure the extent to which the gas deviates from ideal behavior. The van der Waals equation (11.42) also approximates the ideal gas law for sufficiently large v. The van der Waals and virial equations of state may be understood—even derived—if we interpret pressure as a measure of the total energy density of the gas.

The ideal gas equation evolved from a model in which collisions were ignored. As a consequence, the atoms possessed kinetic energy only. Suppose now that we turn on the interatomic forces. These forces do work and convert some of the kinetic energy of colliding atoms into electrical potential energy. The potential energy shared by the atoms is stored only briefly—for the duration of the collision. As the atoms recede from one another their potential energy is converted back into kinetic energy. As a consequence of these collisions a certain amount of energy is stored in the form of electrical potential energy. Because the energy of the gas atoms is not wholly kinetic we expect that any improvement on the ideal gas model must somehow take account of the potential energy. For the ideal gas we found that the pressure was directly proportional to the internal energy density, namely,

$$P = \frac{2}{3}u = \frac{N}{V}kT \tag{11.36'}$$

the internal energy consisting entirely of the kinetic energy of the atoms. The internal energy of a real gas is the sum of the potential and kinetic energies of

the atoms. We therefore modify (11.36) by writing

$$P = \frac{N}{V}kT + \Phi \tag{11.44}$$

the term Φ being proportional to the potential energy density of the gas. Postulating the form (11.44) for P permits us to deduce both the van der Waals equation and the virial equation. To establish the van der Waals equation (11.42), we argue that

(a) Φ is negative, and
(b) Φ is proportional to $(N/V)^2$.

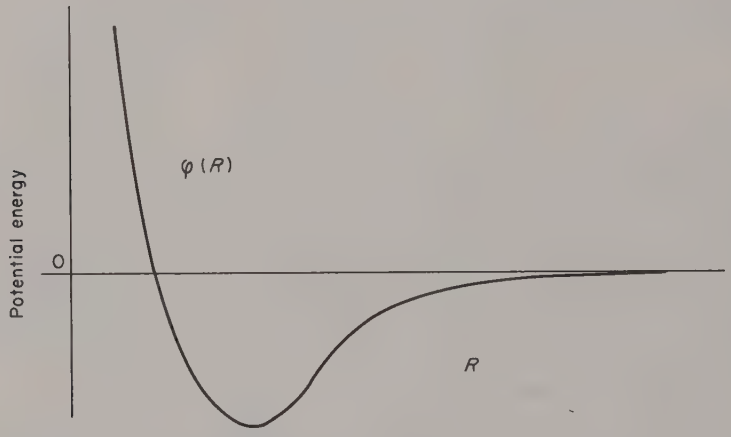


Figure 11.9. Typical form of the interatomic potential. The quantity R denotes the distance between the centers of the two interacting atoms.

To establish (a) we must know how the mutual potential energy of a pair of atoms varies with the distance between them. Figure 11.9 displays the typical form of this potential energy. For large separations (several angstroms) the potential energy is negative and very small in magnitude. The associated force is attractive. Smaller separations result in a more negative potential energy and a stronger attractive force. For sufficiently small separations, the atomic charge clouds overlap to such an extent that the two positively charged nuclei feel each other's presence. When this occurs the force becomes strongly repulsive and the potential energy rises sharply.

The vast majority of encounters between two atoms result in attractive forces for which the potential energy is negative. Only rarely do two atoms collide with such gusto that they attain a positive potential energy—here identified with a repulsive force. Since an overwhelming majority of the collisions involve

a negative potential energy we conclude that the mean potential energy is also negative (a) $\Phi < 0$. To justify (b) we assume that Φ, the interatomic potential energy contribution to the pressure, is proportional to the mean potential energy density of the gas. Let ψ denote the mean potential energy per atom. With N atoms in a volume V we can write the potential energy per unit volume as

$$\Phi = A\frac{N}{V}\psi \tag{11.45}$$

where A is some dimensionless proportionality factor. The question remaining is, "How does ψ depend on N/V?" To answer this we mentally follow a typical atom about and compute its mean potential energy—then ask how the result would be changed if N/V were twice as large. If doubling N/V leaves ψ unchanged, then ψ is independent of N/V. If doubling N/V doubles ψ, then ψ is directly proportional to N/V, and so on. In following the test atom we would discover that most of its collisions are binary—the test atom interacts with only one atom at a time. Occasionally a three-body collision will occur in which the test atom interacts simultaneously with two other atoms. Still rarer are collisions which involve four atoms, and so on. Let us restrict ourselves to consideration of binary collisions for the moment since they are the most frequent type and therefore most important in determining the mean potential energy ψ. By restricting ourselves to binary collisions it is not difficult to show that ψ is directly proportional to N/V. This may be seen as follows: Let N/V be doubled by doubling the number of atoms. With N doubled, the test atom would on the average suffer twice as many binary collisions in any given time interval because the number of targets is twice as great. Doubling the number of collisions doubles the mean potential energy. Reflection shows that doubling N/V by any means—such as halving V with N fixed—has the same effect; the binary collision rate and thus the mean potential energy ψ are doubled. Thus ψ is directly proportional to N/V

$$\psi \propto \frac{N}{V} \tag{11.46}$$

whereupon we conclude that Φ is proportional $(N/V)^2$

$$\Phi = A\frac{N}{V}\psi \propto \left(\frac{N}{V}\right)^2 \tag{11.47}$$

Having established earlier that Φ is negative we can write

$$\Phi = -a\left(\frac{N}{V}\right)^2, \qquad a > 0 \tag{11.48}$$

where a is a positive proportionality factor.

The total pressure $P = (NkT/V) + \Phi$ thus has the form

$$P = \frac{NkT}{V} - \frac{aN^2}{V^2} \tag{11.49}$$

With $v = V/N$ this converts to

$$\left(P + \frac{a}{v^2}\right)v = kT \tag{11.50}$$

which is almost van der Waals' equation of state. We can convert (11.50) into the van der Waals equation (11.42) by remedying one further defect of the ideal gas model. We assumed that the atoms of an ideal gas were mass points. In fact, however, the atoms take up space so that the available volume is less than that of the empty container. Let b denote the average volume excluded by one atom. Then N atoms claim a volume Nb and the effective volume of the container is $V - Nb$. For a rarified gas $V \gg Nb$, that is, the atoms take up only a small fraction of the total volume of the container. Replacing V by $V - Nb$ is equivalent to replacing v by $v - b$, so that (11.50) becomes

$$\left[P + \frac{a}{(v-b)^2}\right](v-b) = kT \tag{11.51}$$

The term $a/(v-b)^2$ already represents a small correction and since $v \gg b$ we can legitimately write

$$\frac{a}{(v-b)^2} \simeq \frac{a}{v^2} \tag{11.52}$$

consistent with the accuracy carried elsewhere. Using this approximation in (11.51) gives the van der Waals equation of state

$$\left(P + \frac{a}{v^2}\right)(v-b) = kT \tag{11.42}$$

Our microscopic approach has yielded the following interpretation of van der Waals' equation of state: The quantity b measures the volume taken up by an individual atom. The term a/v^2 measures the potential energy density and arises because of the predominantly attractive interatomic forces. For given values of v and T, (11.42) shows that the total pressure P is less than it would be in the absence of interatomic forces. This result can be understood from two viewpoints. On the one hand we can think in terms of energy density and postulate that the pressure is proportional to the total energy density. Our arguments then show that the potential energy density contribution to P is negative ($-a/v^2$). We can also understand this reduction in P from a more conventional point of view. Suppose there were no forces between the gas

atoms. There would be a certain pressure on the container walls resulting from atomic impacts. If we now imagine "turning on" the predominantly attractive interatomic forces, it follows that the atoms would not strike the container walls quite so vigorously. The pressure would be somewhat smaller because the attractive forces tend to restrain the motions of the atoms.

The arguments advanced here to obtain the van der Waals equation may be extended to deduce the virial equation. The extension requires inclusion of the effects of interactions between three, four, etc., particles. In general the mean potential energy per atom ψ has the form

$$\psi = \psi_2 + \psi_3 + \psi_4 + \cdots$$

where ψ_s is the contribution arising from the simultaneous interaction of s atoms. Whereas ψ_2 is proportional to N/V, ψ_3 varies as $(N/V)^2$, and ψ_4 as $(N/V)^3$, and so on. Recognition of this point then leads directly to the virial equation of state (11.43).

REFERENCES

A thumbnail history of the kinetic theory of gases will be found in:

E. H. Kennard, "Kinetic Theory of Gases," Chapter 1. McGraw-Hill, New York (1938).

The concept of a degree of freedom and the equipartition of energy theorem are discussed in detail in:

W. Kauzmann, "Kinetic Theory of Gases," Chapter 3. Benjamin, New York (1966).

A general proof of the equipartition theorem for classical systems is given in:

F. C. Andrews, "Equilibrium Statistical Mechanics," Chapter 16. Wiley, New York (1963).

Nonideal gas equations of state are discussed from the microscopic viewpoint in:

W. Kauzmann, "Kinetic Theory of Gases," Chapter 2. Benjamin, New York (1966).

R. D. Present, "Kinetic Theory of Gases," Chapter 6. McGraw-Hill, New York (1958).

F. W. Sears, "An Introduction to Thermodynamics, The Kinetic Theory of Gases, and Statistical Mechanics," 2nd ed., Chapter 11. Addison-Wesley, Reading, Massachusetts (1953).

Chapter 12

Velocity Space Distribution Functions

In this chapter we introduce the concepts of velocity space and velocity space distribution functions. Some of the experimental techniques used to determine velocity distributions are presented in Section 12.6.

12.1 INTRODUCTORY IDEAS

Consider a gas composed of N atoms which occupy a volume V. For simplicity it is assumed that the atoms are identical, for example, all helium atoms or all neon atoms. Further, it is assumed that the gas is in equilibrium with the container walls.

A measuring instrument, being macroscopic by nature, cannot register the effects of a single atom. Any observed meter reading measures a response to the combined action of many atoms. For this reason kinetic theory does not require a detailed microscopic description of the gas. Something less sharp than a specification of the mechanical state will be adequate. Nevertheless, we begin by envisioning a *geometrical* representation of the mechanical state of the gas. The resulting picture suggests a type of description which will meet the needs of kinetic theory. To specify the mechanical state requires that the position

$$\mathbf{r} = (x, y, z)$$

and velocity

$$\mathbf{v} = (v_x, v_y, v_z)$$

be given for each of the N atoms. Geometrically this information may be represented by N points in *position space* and N points in *velocity space*. For a gas of $N \sim 10^{23}$ atoms we can appreciate that there is a high density of points in both position space and velocity space. The high density of such points is important. Many atoms have positions which differ only slightly, and many atoms have velocities which are nearly the same. Also at hand is the experimental fact that instruments respond only to the combined action of many atoms. This suggests that kinetic theory be developed in a fashion which focuses attention on groups of atoms which occupy nearly the same position and which have nearly equal velocities. The key object in such a theory is the particle distribution function. The distribution function concept was introduced in Chapter 10. In general, kinetic theory is concerned with distribution functions which refer to both the positions and the velocities of particles. By dealing with systems which are spatially uniform we abandon half of this conceptual burden—we must "worry" only about the distribution of velocities. From a geometric viewpoint a distribution function measures the density of points, each representing an atom, in a velocity space. Before introducing such distribution functions we digress to develop a familiarity with volume elements in position space and in velocity space.

Consider a volume element (see Figure 12.1)

$$d^3r = dx\,dy\,dz \tag{12.1}$$

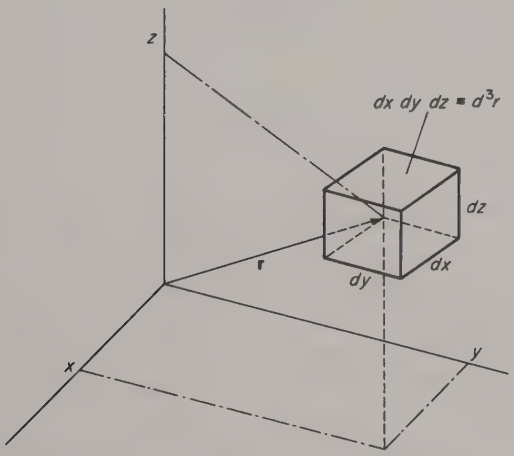


Figure 12.1. Volume element in position space. The volume d^3r is big on a microscopic scale but small by macroscopic standards.

located at

$$\mathbf{r} = (x, y, z)$$

An atom is said to be "in d^3r" if its coordinates lie in the ranges x to $x+dx$, y to $y+dy$, and z to $z+dz$. The volume element d^3r is big on a microscopic scale, but small on a macroscopic scale. Big microscopically means that it contains many atoms—at least most of the time†—thereby allowing a meaningful statistical analysis. Small macroscopically means that measurable properties of the atoms (for example, temperature, pressure) do not vary throughout d^3r. Thus, for example, a meter designed to measure gas pressure at a point records the average effect of many particles over a macroscopic region. Such macroscopic sampling volumes are assumed large by comparison with d^3r.

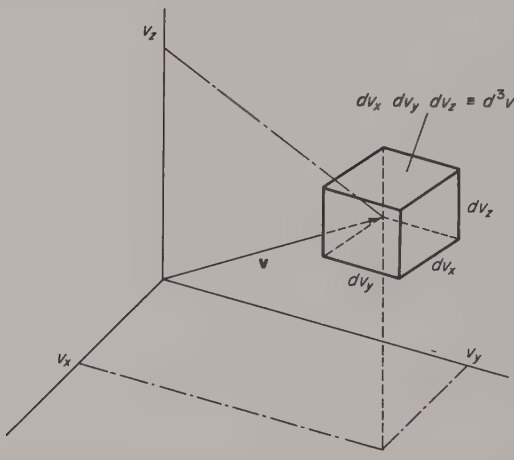


Figure 12.2. Volume element in velocity space. The volume d^3v is big on a microscopic scale but small by macroscopic standards.

At this point we must make a short conceptual leap. Envision first the particles in some volume element d^3r. In general, the particle velocities will be spread over a wide range. We can mentally sort the particles of d^3r into groups having similar velocities. To facilitate this bookkeeping we slice up velocity space in the same fashion we dissected position space. By analogy with d^3r we define a volume element in velocity space by

$$d^3v = dv_x\, dv_y\, dv_z \tag{12.2}$$

† A fluctuation might depopulate the volume element briefly.

In velocity space, d^3v is located at

$$\mathbf{v} = (v_x, v_y, v_z)$$

An atom is said to be "in d^3v" if its velocity vector has components which lie in the range v_x to $v_x + dv_x$, v_y to $v_y + dv_y$, and v_z to $v_z + dv_z$. Figure 12.2, strongly reminiscent of Figure 12.1, shows a typical d^3v. As with d^3r, the volume d^3v is big on a microscopic scale, but small by macroscopic standards. This means that d^3v encompasses a range of velocities large enough to include many atoms. Dealing with volume elements which contain many particles allows us to ignore fluctuations.

12.2 A DISTRIBUTION FUNCTION FOR VELOCITY SPACE

Consider a gas of N atoms in thermodynamic equilibrium. For a system in thermodynamic equilibrium there is no spatial variation in composition—the number of atoms per cubic centimeter is the same throughout the gas. The situation is quite different when we inquire after the distribution of atoms with respect to velocity. Elementary considerations rule out distributions which have every atom moving with the same velocity, or even the same speed. If all atoms had the same speed at some instant, interparticle collisions would quickly result in a spread of atomic speeds. Such distributions cannot represent equilibrium because they are altered by collisions. The equilibrium distribution of velocities is one which is produced and maintained by collisions.

Because the number density is uniform we can ease our conceptual burden by dealing with the full volume V of the gas, rather than with some miniscule chunk of it. If and when it should become necessary to focus our attention on some particular spatial volume element d^3r we may presume that the fraction d^3r/V of all atoms reside therein.

To describe the distribution of molecular velocities we divide velocity space into conveniently shaped and sized elements, named d^3v. Let dN denote the number of molecules in d^3v; we write

$$dN = f(\mathbf{v})\, d^3v \tag{12.3}$$

thereby defining a function $f(\mathbf{v})$ which measures the density of molecules in velocity space. We shall refer to $f(\mathbf{v})$ as the "velocity space distribution function" or, when no confusion can result, simply as the "distribution function." Equation (12.3) is of the utmost importance. You must develop a "seat of the pants" feeling for what it says to be able to master kinetic theory. In arriving at (12.3) we mentally sliced up velocity space. By reversing this procedure—by reassembling the pieces of velocity space and adding the numbers of atoms in each—we arrive at the normalization condition on $f(\mathbf{v})$. This synthesis is

formally accomplished by integrating dN over all segments of velocity space. Thus summing the populations of all velocity space volume elements gives the total number of particles N

$$N = \int_{\mathbf{v}} dN = \int_{\mathbf{v}} f(\mathbf{v})\, d^3v \tag{12.4}$$

The notation $\int_{\mathbf{v}}$ is a symbolic instruction to integrate over all volume elements of velocity space.

....

12.2.1 Explain the nature of $f(\mathbf{v})$ by analogy with the mass density $\rho(\mathbf{r})$.

....

One of the immediate questions is this: "What is the functional form of $f(\mathbf{v})$ for a gas in thermodynamic equilibrium?" The answer was first provided over 100 years ago by James Clerk Maxwell, the man with the equations and the demon! The Maxwellian distribution function for N atoms in thermodynamic equilibrium at a temperature T is

$$f(v) = N\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(\frac{-\frac{1}{2}mv^2}{kT}\right) \tag{12.5}$$

The Maxwellian velocity distribution function is unique—it is the only possible distribution of velocities for a gas in thermodynamic equilibrium.† Two points regarding the Maxwellian distribution should be noted.

(a) The Maxwellian distribution is spherically symmetric in velocity space, that is, $f(v)$ depends on the speed of an atom but not on its direction of motion. This fact will permit us to employ a spherical velocity space in many applications.

(b) The second and more significant feature of the Maxwellian distribution is that the only intrinsic property of the atoms appearing in $f(v)$ is their mass m. Evidently the precise manner in which the atoms interact is immaterial. However, the fact that the atoms do collide is crucially important. Interatomic collisions produce and maintain the Maxwellian distribution. In a very real sense, collisions between atoms drive a gas toward thermodynamic equilibrium and subsequently maintain the equilibrium state.

We have not presented any sort of derivation of the Maxwellian distribution. Such derivations as might be presented now are either cumbersome, or shaky,

† This remark is applicable only within the framework of nonrelativistic, classical kinetic theory. Relativistic and quantum effects modify the Maxwellian form of the equilibrium distribution function.

or both! Statistical mechanics furnishes a rigorous and straightforward derivation of the Maxwellian distribution (compare Chapter 15). Some of the experiments designed to confirm the Maxwellian distribution of velocities are discussed in Section 12.6.

Finally it must be remarked that the uniqueness of the Maxwellian velocity distribution function is limited to the thermodynamic equilibrium domain. Nonequilibrium kinetic theory furnishes numerous examples of non-Maxwellian distribution functions.

12.3 MEAN VALUES

Let $Q(\mathbf{v})$ denote some physical property depending on the velocity of an atom. Prime examples are linear momentum $m\mathbf{v}$ and the kinetic energy $\frac{1}{2}mv^2$. A fundamental hypothesis of kinetic theory is that the measurable quantity corresponding to any such property is Q_{tot}, the sum of the Q's for the N atoms. The $dN(\mathbf{v})$ atoms in the velocity space volume element d^3v each contribute $Q(\mathbf{v})$ to Q_{tot}. The contribution to Q_{tot} from the atoms in d^3v is therefore,

$$Q(\mathbf{v})\,dN(\mathbf{v}) = Q(\mathbf{v})f(\mathbf{v})\,d^3v$$

Integrating over all elements of velocity space gives Q_{tot},

$$Q_{\text{tot}} = \int_{\mathbf{v}} Q(\mathbf{v})f(\mathbf{v})\,d^3v \tag{12.6}$$

Frequently it is desirable to present results on a per particle basis. The mean value per particle $\langle Q \rangle$ is defined by

$$\langle Q \rangle = \frac{Q_{\text{tot}}}{N} \tag{12.7}$$

In terms of $f(\mathbf{v})$ we have

$$Q = \frac{1}{\mathrm{N}} \int_{\mathbf{v}} Q(\mathbf{v})f(\mathbf{v})\,d^3v \tag{12.8}$$

Example 12.1 Total Momentum of a Selected Group of Molecules

Let us determine the total momentum of the molecules in a gas whose velocities are in the positive x direction, assuming they have a Maxwellian distribution. In this case

$$Q = mv_x$$

and Q_{tot} is to be computed for those molecules with positive values of v_x. The

velocity space integration is over all values of v_y and v_z. Thus using (12.6)

$$\mathcal{Q}_{\text{tot}} = \int_{v_x=0}^{+\infty} \int_{v_y=-\infty}^{+\infty} \int_{v_z=-\infty}^{+\infty} mv_x f(v)\, dv_x\, dv_y\, dv_z$$

Inserting the Maxwellian form of $f(v)$ given by (12.5) and noting that

$$v^2 = v_x^2 + v_y^2 + v_z^2$$

gives (the v_y and v_z integrations are identical),

$$\mathcal{Q}_{\text{tot}} = N\left(\frac{m}{2\pi kT}\right)^{3/2} \int_0^{\infty} mv_x \exp\left(-\frac{1}{2}\frac{mv_x^2}{kT}\right) dv_x \left[\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\frac{mv_y^2}{kT}\right) dv_y\right]^2$$

Appendix A furnishes expressions for the two integrals, yielding

$$\mathcal{Q}_{\text{tot}} = \frac{1}{2} N \sqrt{\frac{2mkT}{\pi}}$$

Half of the N molecules have positive values of v_x, so that the quantity $\sqrt{2mkT/\pi}$ is the mean momentum per particle in the positive x direction. ✧

....

12.3.1 (a) Determine (by symmetry considerations, preferably) the mean velocity $\langle \mathbf{v} \rangle$ for a gas whose atoms have a Maxwellian distribution of velocities.

(b) The flow of heat in a gas is directly proportional to the mean value of the quantity $\frac{1}{2}mv^2\mathbf{v}$. Determine $\langle \frac{1}{2}mv^2\mathbf{v} \rangle$ (again by symmetry considerations, preferably) for a gas whose atoms have a Maxwellian distribution of velocities. Are your conclusions in accord with the fact that the Maxwellian distribution describes a gas in thermodynamic equilibrium?

....

We have previously noted that the Maxwellian distribution function is spherically symmetric in velocity space. That is, the Maxwellian form of $f(\mathbf{v})$ depends on the magnitude of $\mathbf{v}$ but not on its direction. It will be to our everlasting advantage to momentarily turn aside from kinetic theory proper and study spherical coordinate and velocity spaces.

12.4 SPHERICAL COORDINATE SPACE AND SPHERICAL VELOCITY SPACE

Frequently the symmetries inherent in a physical situation dictate the use of coordinates other than the familiar rectangular system. Of particular importance is the system of spherical coordinates and the associated elements of area

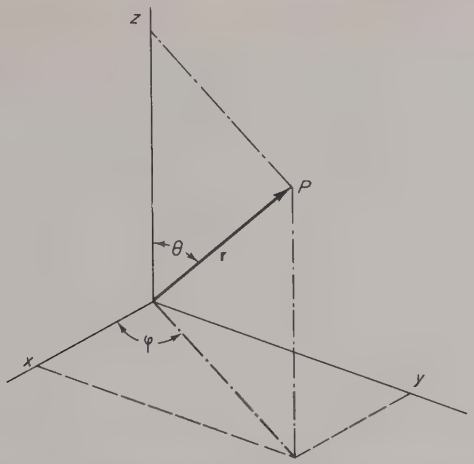


Figure 12.3. Spherical coordinates. The relationships between the spherical coordinates (r, θ, φ) and the Cartesian coordinates (x, y, z) are given by (12.9).

and volume. Figure 12.3 shows the relationships between the Cartesian and spherical coordinates of a point P. The position vector $\mathbf{r}$ directed from the origin to P may be specified by giving the values of its Cartesian components x, y, and z. An equivalent specification is rendered in spherical coordinates by giving the values of r, θ, and φ. From Figure 12.3 we see that the two sets of coordinates are related by

$$\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{aligned} \tag{12.9}$$

and that

$$r^2 = x^2 + y^2 + z^2$$

A differential element of area in spherical coordinates is generated by sweeping the radius vector along two mutually perpendicular directions: First through an angle $d\theta$, from θ to $\theta + d\theta$ with φ held constant, and second through an angle $d\varphi$, from φ to $\varphi + d\varphi$ with θ held constant. As Figure 12.4 illustrates, these steps generate two sides of the differential element of area

$$dA = r^2 \sin\theta \, d\theta \, d\varphi$$

Note that an integration over the surface of a sphere is effected by integrating from $\varphi = 0$ to $\varphi = 2\pi$ and from $\theta = 0$ to $\theta = \pi$. For example, the total surface area of a sphere is given by

$$A_{\text{sph}} = r^2 \int_{\theta=0}^{\pi} \sin\theta \, d\theta \int_{\varphi=0}^{2\pi} d\varphi = 4\pi r^2$$

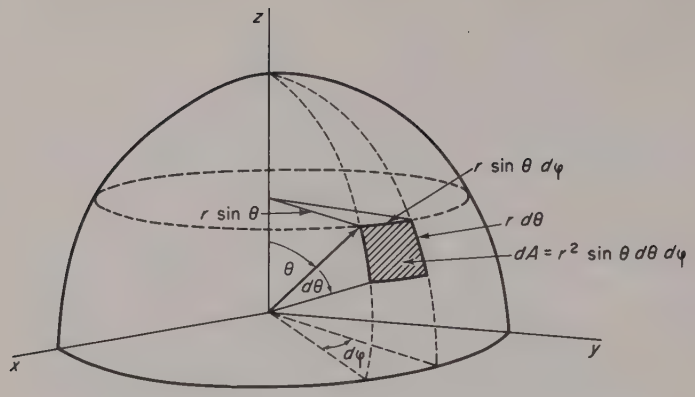


Figure 12.4. Differential element of area in spherical coordinates.

The volume element in spherical coordinates is obtained by multiplying the surface area element by dr. Thus†

$$d^3r = dA\,dr = r^2\,dr\sin\theta\,d\theta\,d\varphi \tag{12.10}$$

Figure 12.5 shows d^3r. The volume of a spherical shell is

$$A_{\text{sph}}\,dr = 4\pi r^2\,dr \tag{12.11}$$

obtained by integrating d^3r over the surface of the sphere, that is, over θ and φ.

The use of spherical coordinates is readily extended to velocity space. In place of the rectangular components of $\mathbf{v}$

$$\mathbf{v} = (v_x, v_y, v_z)$$

we can use the spherical components

$$\mathbf{v} = (v, \theta, \varphi)$$

The velocity space volume element becomes

$$d^3v = v^2\,dv\sin\theta\,d\theta\,d\varphi \tag{12.12}$$

The pictures are the same as those of Figures 12.3–12.5 with x replaced by v_x, r replaced by the speed v, and so on. The volume of a spherical shell of radius v and thickness dv is

$$v^2\,dv\int_{\theta=0}^{\pi}\sin\theta\,d\theta\int_{\varphi=0}^{2\pi}d\varphi = 4\pi v^2\,dv \tag{12.13}$$

---

† The symbolism d^3r is a generic one used to designate three-dimensional volume elements. It is not restricted to indicate the rectangular volume element $dx\,dy\,dz$.

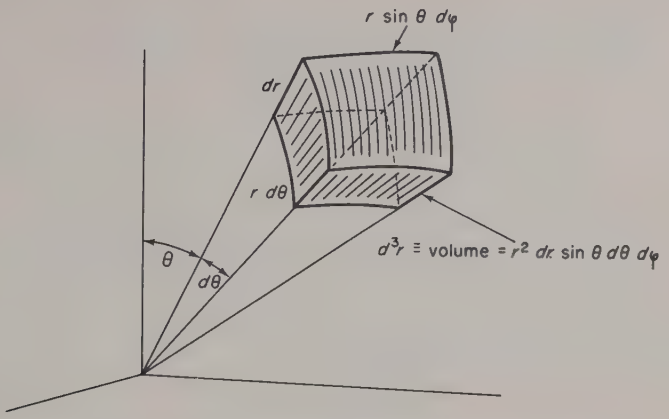


Figure 12.5. Volume element for spherical coordinates.

When the symmetry of velocity space dictates its use, we may take $d^3v = 4\pi v^2\, dv$ and integrate over the speed alone.

We shall make liberal use of spherical velocity space throughout the remainder of our work in kinetic theory.

Example 12.2 Mean Kinetic Energy for the Maxwellian Distribution

As an illustration of the use of spherical velocity space we compute the mean kinetic energy in a gas by using (12.8) with $Q = \frac{1}{2}mv^2$. We do this using the Maxwellian velocity space distribution function, given by (12.5). Because both $Q(v) = \frac{1}{2}mv^2$ and the distribution function are spherically symmetric we can integrate over spherical shells of volume $d^3v = 4\pi v^2\, dv$. Thus

$$\left\langle \frac{1}{2}mv^2 \right\rangle = \frac{1}{\mathrm{N}} \int_{v=0}^{\infty} \frac{1}{2}mv^2 f(v)\, 4\pi v^2\, dv$$

Inserting the Maxwellian form of $f(v)$ and introducing the dimensionless variable $u^2 = \frac{1}{2}mv^2/kT$ converts this to

$$\left\langle \frac{1}{2}mv^2 \right\rangle = \frac{4kT}{\sqrt{\pi}} \int_0^{\infty} u^4 \exp(-u^2)\, du$$

The Gaussian integral has the value $\frac{3}{8}\sqrt{\pi}$ (see Appendix A), which leaves the result

$$\langle \tfrac{1}{2}mv^2 \rangle = \tfrac{3}{2}kT$$

The total kinetic energy is $K = N\langle \frac{1}{2}mv^2 \rangle = \frac{3}{2}NkT$. It should be noted [compare Equation (11.23)] that this is the result we were obliged to assume in our derivation of the ideal gas law. ✧

.....

12.4.1 Assuming a Maxwellian distribution of velocities, evaluate the standard deviation of the kinetic energy,

$$\sigma(K) = \sqrt{\langle K^2 \rangle - \langle K \rangle^2}$$

.....

12.5 THE DISTRIBUTION OF MOLECULAR SPEEDS

For distribution functions which are spherically symmetric in velocity space, such as the Maxwellian, the quantity

$$4\pi v^2 f(v)\, dv = dN \tag{12.14}$$

gives the number of atoms with speeds in the range from v to $v+dv$. Thus the quantity $4\pi v^2 f(v)$ gives the distribution of atomic speeds. Figure 12.6 is a plot of $4\pi v^2 f(v)$ versus v for the Maxwellian distribution. Using the Maxwellian form of $f(v)$ gives

$$4\pi v^2 f(v) = 4\pi N \left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp\left(\frac{-mv^2}{2kT}\right) \tag{12.15}$$

The most immediate advantage of dealing with the distribution of speeds is evidenced by Figure 12.6 itself. The distribution of speeds is readily visualized and plotted. This promotes other advantages. For example, if we ask for the most probable atomic speed, a glance at Figure 12.6 reveals that this is the speed for which $4\pi v^2 f(v)$ is a maximum. The equation defining the most

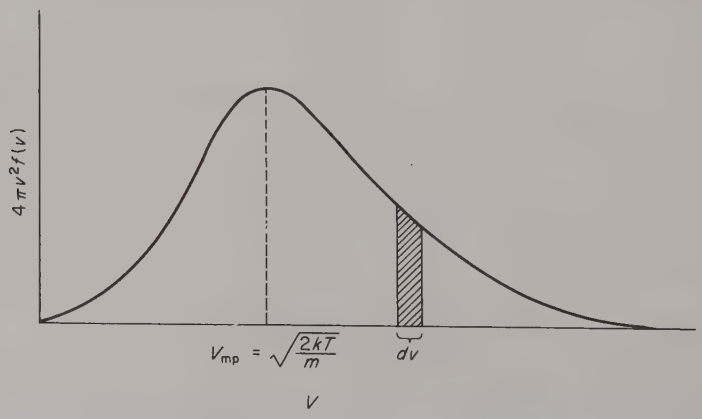


Figure 12.6. The Maxwellian distribution of atomic speeds. Shaded area $= 4\pi v^2 f(v)\, dv$.

probable speed is thus

$$\frac{d}{dv}(4\pi v^2 f(v)) = 0$$

For the Maxwellian distribution we obtain (A = constant)

$$\frac{d}{dv} Av^2 \exp\left(\frac{-\frac{1}{2}mv^2}{kT}\right) = 2Av\left(1 - \frac{mv^2}{2kT}\right)\exp\left(\frac{-\frac{1}{2}mv^2}{kT}\right) = 0$$

The equation determining the most probable speed v_{mp} is thus

$$2Av\left(1 - \frac{mv^2}{2kT}\right)\exp\left(\frac{-\frac{1}{2}mv^2}{kT}\right) = 0$$

This yields

$$v_{mp} = \sqrt{\frac{2kT}{m}} \tag{12.16}$$

Example 12.3 Mean Kinetic Energy for an Electron Gas

Metals such as silver and copper are excellent conductors of electricity. Their high electrical conductivity stems from the fact that each atom in the atomic lattice is readily "coaxed" into volunteering an electron for the task of carrying current. The conducting material can be regarded as a gas of electrons confined within a container-like lattice of atoms and positively charged ions. If the electrons are assumed to have a Maxwellian distribution of speeds, grave troubles arise (see Chapter 17). Instead one finds that $f(v)$ is essentially constant from $v = 0$ up to a characteristic speed $v = v_F$, the so-called Fermi speed. Beyond v_F, $f(v)$ is zero. Thus

$$f(v) = \begin{cases} A, & 0 \leqslant v \leqslant v_F \\ 0, & v > v_F \end{cases} \tag{12.17}$$

The value of A is fixed by the normalization condition

$$N = \int_{\text{all speeds}} dN = \int_0^\infty 4\pi v^2 f(v)\, dv$$

This establishes the relationship of N, v_F, and A

$$N = \frac{4\pi}{3} v_F{}^3 A \tag{12.18}$$

Let us compute the mean kinetic energy of an electron, and express it in terms of the Fermi energy

$$\varepsilon_F \equiv \tfrac{1}{2} m v_F{}^2$$

which is the maximum energy which any electron can have according to the distribution (12.17). We have

$$\langle K \rangle = \frac{1}{N} \int \tfrac{1}{2} m v^2 \, dN = \frac{1}{N} \int_0^{v_F} \tfrac{1}{2} m v^2 \, 4\pi v^2 A \, dv$$

Using (12.18) to express A/N in terms of v_F gives

$$\langle K \rangle = \tfrac{3}{5}(\tfrac{1}{2} m v_F^{\,2}) = \tfrac{3}{5}\varepsilon_F \tag{12.19}$$

Thus the mean kinetic energy is a very healthy fraction of the maximum energy ε_F. ✧

....

12.5.1 Sketch $4\pi v^2 f(v)$ for the electron gas distribution function given by Equation (12.17).

(a) What is the mean electron speed?
(b) What is the most probable electron speed?

The Fermi speed v_F is related to the number density of electrons n_e by the equation

$$v_F = \frac{h}{m}\left(\frac{3n_e}{8\pi}\right)^{1/3}$$

where $h = 6.624 \times 10^{-27}$ erg-sec is Planck's constant and $m = 0.91 \times 10^{-27}$ gm is the electron mass. Assume that each atom in metallic silver donates one electron, so that n_e is equal to the number of silver atoms per unit volume. The density of silver at room temperature is 10.5 gm/cm^3 and the mass of a single silver atom is approximately 1.8×10^{-22} gm.

(c) Compute v_F and ε_F. How does v_F compare with the speed of sound? How does ε_F compare with the characteristic thermal energy $\frac{3}{2}kT$, at $T = 300$ K? Would you describe the electron gas as hot or cold by comparison with its surroundings?

12.5.2 Compute the average speed of molecules at $T = 300$ K in

(a) He
(b) N_2

assuming a Maxwellian distribution of velocities.

(c) The velocity of sound is given by

$$v_s = \sqrt{\frac{\gamma kT}{m}} = \sqrt{\frac{\gamma RT}{M}}$$

where γ is the ratio of heat capacities ($\gamma = C_p/C_v$). Using the ideal gas values of γ for monatomic and diatomic gases determine the ratio of the speeds of sound in helium and nitrogen (both gases at same temperature). The relation

$$f\lambda = v_s$$

relates the frequency f, wavelength λ, and speed of sound. The human "audio system" may be thought of as an open-ended resonant chamber in which the wavelengths of sound waves which can propagate are fixed by the geometry of the throat and mouth. There is a maximum wavelength which can propagate. To this maximum wavelength there corresponds a minimum frequency called the fundamental frequency. Lower frequencies (longer wavelengths) are severely attenuated. They are cut off and do not propagate. The fundamental frequency of the adult male is about 100 Hz. (1 Hz = 1 cycle/sec), and determines the pitch of the voice. The waves propagating in the throat and mouth are a superposition of waves of the fundamental frequency and integral multiples thereof (overtones, harmonics, partials).

(d) If the lungs are filled with helium the resulting speech has a high-pitched character (demonstration recommended). Explain this and estimate the minimum frequency emanating from the mouth.

....

12.6 EXPERIMENTAL VERIFICATION OF THE MAXWELLIAN DISTRIBUTION

A number of investigators have verified the Maxwellian distribution of speeds. We discuss two of these efforts; the novel gravitational deflection study by Estermann, Simpson, and Stern (1947) and the more recent and precise experiment by Miller and Kusch (1955). A review and comprehensive bibliography of other investigations is given by Ramsey.

Figure 12.7 shows a schematic of the apparatus used by Estermann, Simpson, and Stern. A small sample of cesium in an oven heated to about 450 K provided a source of atoms. The oven slit O and a collimating slit S produced a narrow beam of atoms. The detector D was a hot tungsten wire. A

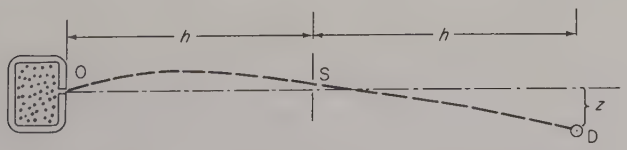


Figure 12.7. Schematic for velocity distribution measurements.

cesium atom striking the wire would be ionized and the resulting positive cesium ions collected by a negatively charged cylinder surrounding the wire. The resulting ion current was thus directly proportional to the number of cesium atoms which reached the detector. The vertical position of the detector could be varied. The slit S was positioned midway between the oven slit and the detector. An atom leaving the oven with the speed v would fall a distance $z = gh^2/v^2$ below the level of the slits, with g denoting the acceleration of gravity and $2h$ the oven-to-detector distance. The distribution of atomic velocities in the beam could thus be inferred by measuring the relative numbers (relative ion currents) for different values of z. The experimental results confirmed the Maxwellian distribution law although there were significant discrepancies for the low speed atoms resulting from scattering of the beam at the slits.†

....

12.6.1 With reference to Figure 12.7:

(a) Show that the distance of fall is $z = gh^2/v^2$. Note that an atom which reaches the detector must leave the oven along a slightly elevated trajectory such that it falls back to the same level in travelling the distance h. Thus the atoms are travelling downward as they pass through the slit S.

(b) Compute the distance z for a cesium atom which travels with the most probable atomic speed ($\sqrt{2kT/m}$) for a temperature of 450 K. The distance h was 100 cm in the experiments of Estermann, Simpson, and Stern.

(c) The slit and the detector wire had widths of 0.02 mm. How would these finite widths affect the measurements?

....

Miller and Kusch used a long rotating cylinder (see Figure 12.8) with a large number of helical grooves cut along its surface. The angular pitch of the grooves was φ_0. The cylinder was positioned so as to act as a shutter in the path of an atomic beam. Only atoms with particular speeds could pass through the grooves to the detector beyond. When the cylinder rotates with an angular velocity ω, atoms with speed

$$v_0 = \frac{\omega L}{\varphi_0} \tag{12.20}$$

† The distribution of speeds in the oven is presumably Maxwellian. The distribution of atomic speeds in the beam differs from that in the oven. The faster moving atoms are more likely to escape the oven with the result that the distribution of speeds in the beam is proportional to the product of the distribution function and the speed v. This point is discussed further in Chapter 13, after the concept of particle flux has been developed (see Problem 13.3.1).

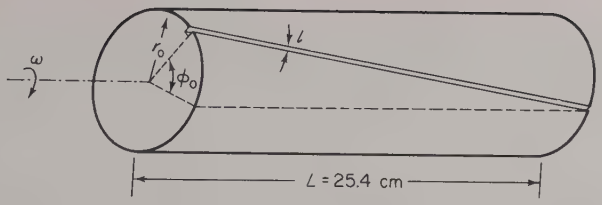


Figure 12.8. Rotating cylinder with grooves used as a velocity selector. L = 25.4 cm; l = width of groove.

can pass without striking the walls of the groove. By varying the angular velocity ω the relative distribution of atomic speeds in the beam could be ascertained. The experiment gave excellent agreement with the expected Maxwellian distribution law for beams of thallium and potassium atoms emanating from an oven operated at temperatures ranging from 466 K to 944 K.

....

12.6.2 Derive Equation (12.20). Note that if the atom is to pass through the groove without striking the walls, it must travel the length of the cylinder L in the same time required for the cylinder to rotate through the angle φ_0.

12.6.3 The finite width l of the grooves in the Miller–Kusch apparatus allowed atoms with a narrow range of speeds to pass through the velocity selector for a fixed angular speed. The limiting values of the speeds correspond to atoms which enter a groove at one side and leave at the opposite side. If r_0 is the radius of the cylinder, show that the maximum and minimum speeds which pass are

$$v_{\max} = \frac{v_0}{1-\gamma}, \qquad v_{\min} = \frac{v_0}{1+\gamma}$$

where $\gamma = l/r_0\varphi_0$. Compute the velocity spread $\Delta v = v_{\max} - v_{\min}$ for $v_0 = 6.3 \times 10^4$ cm/sec, $l = 0.0424$ cm, $r_0 = 10.00$ cm, $\varphi_0 = 2\pi/74.7$ rad.

....

REFERENCES

The concept of velocity space is basic to the development of kinetic theory. See, for example:

E. H. Kennard, "Kinetic Theory of Gases," Chapter 2. McGraw-Hill, New York (1938).

The distribution function concept and its relation to mean values is developed in the reference above as well as in:

A. F. Brown, "Statistical Physics," Chapter 3. Edinburgh Univ. Press, Edinburgh (1968).

Two experiments which verify the Maxwellian distribution are reported in:

I. Esterman, O. C. Simpson, and O. Stern, The Free Fall of Atoms and the Measurement of the Velocity Distribution in a Molecular Beam of Cesium Atoms, *Phys. Rev.* **71**, 238 (1947).

and

R. C. Miller and P. Kusch, Velocity Distribution in Potassium and Thallium Atomic Beams, *Phys. Rev.* **99**, 1314 (1955).

Detailed accounts of velocity distribution measurements are given in:

N. F. Ramsey, "Molecular Beams," Chapter 2. Oxford Univ. Press, London (1956).

Chapter 13

Transport Processes

13.1 INTRODUCTION

Emphasis was laid previously on the fact that real processes are irreversible. A distinguishing characteristic of irreversible processes is the transport of molecular properties. Molecular properties of prime importance include mass, momentum, and energy. The molecules of a gas, in moving about, carry or transport these molecular properties. If the gas is in equilibrium, there are no net flows of mass, momentum, or energy. That is, if an observer could watch the molecules enter and leave some small volume element, he would find that equal quantities of mass enter and leave, so that no net change occurs. The inward and outward flows of energy and momentum are similarly balanced. A nonequilibrium situation results in a net flow of one or more molecular properties. For example, diffusion results in a net flow of particles. Momentum transport is responsible for the viscosity (fluid friction) of gases and liquids, and heat conduction results in the transport of internal energy.

The basic flow concept is that of a flux. The fluxes of mass, momentum, and energy are illustrated in Section 13.2. In Section 13.3 we use the velocity space distribution function to determine the flux of molecules through a leak, that is, the rate at which a confined gas escapes through a small hole in the container wall. In Section 13.4 the flux of internal energy is related to the process of heat conduction.

13.2 FLUXES

The concept of a particle flux is basic to the discussion of transport phenomena. Consider a group of particles, all of which move with the same velocity $\mathbf{v}_0$. Let n denote the number density of the particles, the number per unit volume,†

$$n = \frac{N}{V}$$

The particle flux is a vector quantity defined by

$$\mathbf{S} = n\mathbf{v}_0 \tag{13.1}$$

To establish the physical significance of the flux, imagine that the particles are moving along a cylindrical pipe of cross-sectional area A. The number of particles passing any cross section in time t is

$$nAv_0 t = SAt$$

This result may be established as follows: Particles which pass a given cross section during a time t lie somewhere within a cylinder of length $v_0 t$ and thus of volume $Av_0 t$ (see Figure 13.1). As the number density is n, the number of

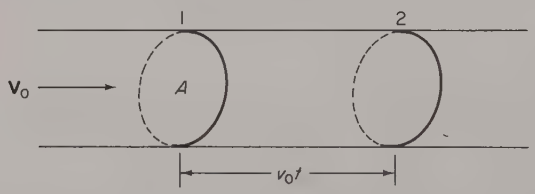


Figure 13.1. Particles which pass point 1 during a time interval t lie somewhere within the cylindrical volume $Av_0 t$.

particles in the volume $Av_0 t$ is $nAv_0 t = SAt$. The particles in this volume are those which entered in time t; particles there originally have moved out. Thus the number of particles which cross area A in time t is given by SAt, with S being the particle flux. This interpretation requires that the area A be oriented perpendicular to the flow of particles. Had we chosen a different surface (see Figure 13.2) of area A' we would have concluded that the number crossing in time t is

$$SA' \cos\theta t = SAt, \qquad A' \cos\theta = A$$

The area $A = A' \cos\theta$ is the projection of A' normal to the particle flow.

† Note that this use of n differs from an earlier usage where the same symbol denoted the number of moles.

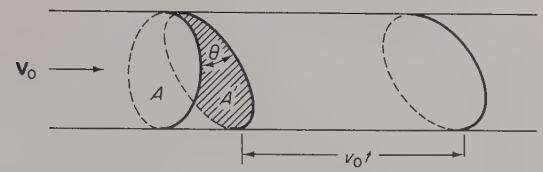


Figure 13.2. It is the projection of A' normal to the flow of particles ($A' \cos\theta$) which determines the number of particles crossing A' in time t. The volume of the slanted cylinder is $A' \cos\theta\, v_0 t = A v_0 t$.

If SAt gives the number crossing A in time t, we can interpret the particle flux as the number of particles that cross per unit area per unit time. The flux S is numerically (but not dimensionally) equal to the number of particles which cross unit area in unit time. The flux is an intensive quantity characterizing the flow of molecules, whereas the total number of particles passing through the area A in time t is an extensive quantity, being directly proportional to both A and t. It also follows that the rate at which particles cross an area A normal to the flow is SA.

$$SA = nv_0 A = \frac{\text{number of particles crossing the normal area } A}{\text{sec}}$$

Example 13.1 Molecular Flux

A pipe transports 60 gal of water/min. The cross-sectional area of the pipe is 40 cm^2. What is the flux of water molecules? We must first determine the number of molecules in 1 gal of water. Recalling that noble Anglo-Saxon refrain, "A pint's a pound the world around," reminds us that 1 pt of water weighs 1 lb so that 1 gal weighs 8 lb. A weight of 1 lb has a mass of approximately 450 gm. Therefore,

$$(\text{mass of 1 gal of water}) \simeq 8 \text{ lb} \frac{450 \text{ gm}}{\text{lb}} = 3600 \text{ gm}$$

The mass of 1 mole of water is 18 gm. Thus

$$(\text{number of molecules in 1 gal of water})$$

$$\simeq 3600 \frac{\text{gm}}{\text{gal}} \frac{1 \text{ mole}}{18 \text{ gm}} 6 \times 10^{23} \frac{\text{molecules}}{\text{mole}}$$

$$= 1.20 \times 10^{26} \frac{\text{molecules}}{\text{gal}}$$

Because 60 gal/min is 1 gal/sec, it follows that 1.2×10^{26} molecules flow across 40 cm^2 each sec. The flux is therefore 3×10^{24} molecules/sec. ✧

....

13.2.1 The density of the water in Example 13.1 is 1 gm/cm^3. Show that the speed of the flow is approximately 90 cm/sec.

13.2.2 A sports arena has a capacity of 13,200. It takes 15 min to empty a capacity crowd using 24 doors. Each door has a usable emptying area of 15 sq. ft. What is the people flux when a capacity crowd exits as specified?

....

We must extend the notion of a particle flux to the microscopic domain. If $f(\mathbf{v})$ is the distribution function, the number of atoms with velocities in d^3v is $dN = f(\mathbf{v})\,d^3v$ and the corresponding particle flux is $\mathbf{v}\,dN/V$

$$\mathbf{dS} = \mathbf{v}f(\mathbf{v})\frac{d^3v}{V} \tag{13.2}$$

The net particle flux is found by integrating over the full range of velocities

$$\mathbf{S} = \int_{\substack{\text{all}\\ \text{veloc.}}} \mathbf{dS} = \int_{\mathbf{v}} \mathbf{v}f(\mathbf{v})\frac{d^3v}{V} \tag{13.3}$$

It is frequently desirable to introduce other fluxes. For example, the mass flux is obtained by multiplying the particle flux by the mass per particle. This gives the directed rate of mass flow per unit area. The portion of the mass flux contributed by atoms in d^3v is

$$\mathbf{dM} \equiv m\,\mathbf{dS} = m\mathbf{v}f(\mathbf{v})\frac{d^3v}{V}$$

where m is the mass of a single atom. The total mass flux is

$$\mathbf{M} = \int_{\substack{\text{all}\\ \text{veloc.}}} \mathbf{dM} = \frac{1}{V}\int_{\mathbf{v}} m\mathbf{v}f(\mathbf{v})\,d^3v = m\mathbf{S}$$

The mass flux is evidently the rate at which mass flows per unit area. In general, if $\Phi(\mathbf{v})$ represents some molecular property, the corresponding flux of the property resulting from the particles in d^3v is

$$\mathbf{d}\mathscr{F} \equiv \Phi(\mathbf{v})\,\mathbf{dS} = \Phi(\mathbf{v})\,\mathbf{v}f(\mathbf{v})\frac{d^3v}{V}$$

and the total flux of the quantity is

$$\mathscr{F} = \frac{1}{V}\int_{\substack{\text{all}\\ \text{veloc.}}} \Phi(\mathbf{v})\,\mathbf{v}f(\mathbf{v})\,d^3v \tag{13.4}$$

The fluxes of mass, momentum, and energy are of primary importance in the kinetic theory of gases. In fact, the hydrodynamic equations for a gas may be derived from consideration of just these three fluxes. In a plasma (an assembly of ionized particles) the flux of electric charge also must be considered. By combining the equations of kinetic theory with Maxwell's equations one can develop magnetohydrodynamics, the hydrodynamics of electrically conducting fluids.

Example 13.2 Kinetic Energy Flux

A hailstorm produces a vertical flux of hailstones. Each hailstone has the same mass (12 gm) and reaches the ground with the same speed (800 cm/sec). The number density of hailstones is $0.001/\text{cm}^3$. Determine the particle flux **S** and the kinetic energy flux of the hailstones.

The vector **S** is directed vertically downward. Its magnitude is

$$S = nv_0 = \left(\frac{0.001}{\text{cm}^3}\right)\left(800\ \frac{\text{cm}}{\text{sec}}\right) = \frac{0.8}{\text{cm}^2\text{-sec}}$$

The flux of kinetic energy is also downward. Since each hailstone has the same mass and speed the flux of kinetic energy $\mathscr{F}_k$ is the kinetic energy of one stone times the particle flux

$$\mathscr{F}_k = \tfrac{1}{2}mv_0^2 S = 3.072 \times 10^6\ \text{ergs/cm}^2\text{-sec} \quad \diamondsuit$$

....

13.2.3 Determine the fluxes of

(a) mass,
(b) linear momentum,

for the hailstones of Example 13.2.

....

13.3 MOLECULAR EFFUSION

The escape of gas through a small hole in a container is called *effusion*. In general, gas atoms proceed in both directions through such a hole. The net rate of effusion depends on the difference of the particle fluxes into and out of the container. For simplicity we assume that a vacuum surrounds the container so that only the escape of molecules need be considered. We also assume that the hole is large in comparison to atomic sizes. An atom, upon arriving at the opening, is able to pass through freely. On the other hand the hole must not be so large that the outward flow materially alters the distribution of velocities

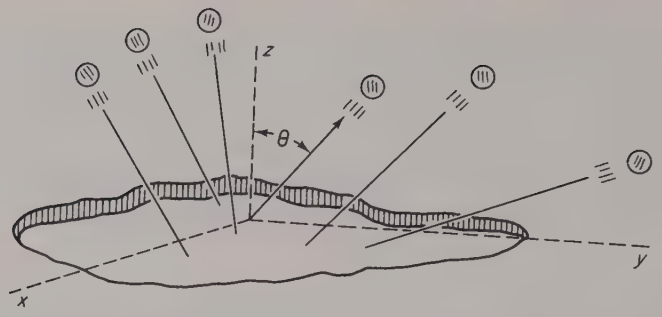


Figure 13.3. Geometry for computing the rate of molecular effusion. The polar angle θ is measured from the direction of the positive z axis. Only molecules with θ ranging from 0 to $\pi/2$ are moving outward.

of the remaining atoms. Collisions between atoms will be sufficiently numerous to maintain the form of the velocity distribution provided the dimensions of the hole do not exceed the mean free path length. For the moment we assume only that the distribution function is spatially uniform and spherically symmetric in velocity space. Later the Maxwellian distribution is called forth to provide a numerical estimate of the rate of effusion.

Figure 13.3 displays the geometry to be used. The positive z axis is directed out of the containing vessel. The opening lies in the x–y plane. The spherical symmetry of the velocity distribution guarantees that the x and y components of the flux are zero. Furthermore, only atoms with positive z components of velocity can escape. From (13.3) it follows that the flux is

$$S_z = \frac{1}{V} \int_{\substack{\text{all veloc.} \\ \text{with } v_z > 0}} v_z f(v)\, d^3v \tag{13.3a}$$

Spherical velocity space variables are called for

$$v_z = v \cos\theta$$

$$d^3v = v^2\, dv \sin\theta\, d\theta\, d\varphi$$

We are concerned only with atoms moving in the positive z direction. Consequently the integration over the polar angle θ ranges from 0 to $\pi/2$. The azimuth angle φ runs from 0 to 2π as usual. With these provisions

$$S_z = \frac{1}{V} \int_0^\infty v f(v)\, v^2\, dv \int_0^{\pi/2} \cos\theta \sin\theta\, d\theta \int_0^{2\pi} d\varphi$$

The angular integrations result in the factor π

$$S_z = \pi \int_0^\infty v f(v)\, v^2 \frac{dv}{V} = \frac{N}{4V} \frac{\int_0^\infty v f(v)\, 4\pi v^2\, dv}{N}$$

Noting that $N/V = n$, the number density, and that $(1/N)\int_0^\infty vf(v)\,4\pi v^2\,dv = \langle v\rangle$ is the mean atomic speed leaves us with

$$S_z = \tfrac{1}{4}n\langle v\rangle \tag{13.5}$$

Experiments by Knudsen in 1909 indirectly verified this expression for the effusive flux.

Recalling that $S = nv_0$ for a group of atoms with the same velocity, one more or less expects the $n\langle v\rangle$ dependence of S_z. Many aspects of such processes may be understood qualitatively by assuming that each atom moves with the mean speed or that each possesses the average energy. In this vein, it is helpful to introduce the idea of an average atom—one which moves with the mean speed or which possesses the average energy. In the next section we use the notion of the average atom to discuss thermal conduction. A rigorous treatment of transport phenomena in gases requires the use of the Boltzmann equation, a somewhat fearsome integro-differential equation.

....

13.3.1 The average speed in a beam of atoms effusing from an oven at temperature T is greater than the average speed of atoms within the oven, because the faster atoms are more likely to escape. Show that the average speed of an atom in the beam is $3\pi/8$ greater than the average speed of an atom in the oven, assuming a Maxwellian distribution of speeds in the oven. [*Hint:* The probability that an atom of speed v will escape from the oven is directly proportional to its speed. It is also proportional to the probability that there is an atom of speed v in the oven. In other words, the probability of finding an atom in the beam with its speed in the range v to $v+dv$ is a conditional probability.]

....

13.4 THERMAL CONDUCTION

We may understand the phenomenon of thermal conduction in a gas in terms of a flux of internal energy. For simplicity we assume that the temperature of the gas varies only along the z direction. Experimentally it is observed that the heat flux (heat flow per unit area per sec) is proportional to the temperature gradient $\partial T/\partial z$. With q_z denoting the heat flux we may write

$$q_z = -\kappa\frac{\partial T}{\partial z} \tag{13.6}$$

The constant κ is called the *thermal conductivity*. The minus sign expresses the fact that heat flows from a higher to a lower temperature. Equation (13.6) is Fourier's law of heat conduction.

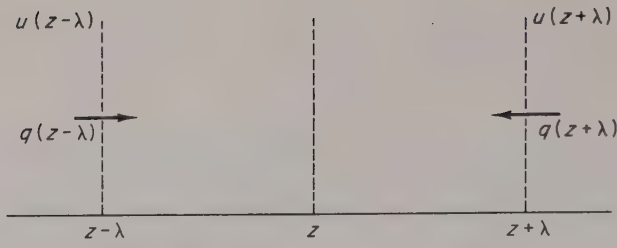


Figure 13.4. The average atom crossing the plane $z =$ constant originates from a plane at $z+\lambda$ or at $z-\lambda$.

....

13.4.1 Can a system in thermodynamic equilibrium conduct heat? Explain.

....

We can deduce (13.6) by considering the net flux of energy across an imaginary plane perpendicular to the z axis (see Figure 13.4). On the average, an atom crossing this plane will travel one mean free path λ before suffering a collision. It follows that atoms reaching this plane have traveled a distance λ, on the average, since undergoing their last collision. The general effect of collisions is to randomize the velocities of the colliding particles. On the average then, the momentum and energy of an atom are determined by its most recent collision. We therefore picture the average atoms which reach z as emerging from planes at $z+\lambda$ and at $z-\lambda$ and having energies and momenta appropriate to these planes.

Let u denote the mean internal energy per atom. Now u depends on T and thus indirectly on position z. In order to focus all attention on the thermal conduction we assume that there is no net particle flux, that is, no diffusion. We take the particle flux (directed toward z) to be $\frac{1}{4}n\langle v\rangle$ at both $z+\lambda$ and $z-\lambda$. This gives zero net particle flux at z. The internal energy flux of the average atoms is u times the particle flux

$$q = \frac{n\langle v\rangle}{4}u \tag{13.7}$$

The net internal energy flux at z is the sum of the oppositely directed fluxes emanating at $z\pm\lambda$. We have

$$q(z-\lambda) = \frac{n\langle v\rangle}{4}u(z-\lambda)$$

$$q(z+\lambda) = -\frac{n\langle v\rangle}{4}u(z+\lambda)$$

It is the net flux $q(z+\lambda) + q(z-\lambda)$, which we identify as the heat flux q_z

$$q_z = \frac{n\langle v\rangle}{4}[u(z-\lambda) - u(z+\lambda)] \tag{13.8}$$

Note that there is no net flux of atoms; equal numbers pass z in either direction. Nevertheless a net flow of thermal energy results because of a difference in the internal energy at $z+\lambda$ and at $z-\lambda$. Atoms coming from the region of higher temperature carry the greater energy. A Taylor series expansion for the internal energy gives

$$u(z\pm\lambda) = u(z) \pm \lambda\left(\frac{\partial u}{\partial z}\right) + \cdots \tag{13.9}$$

In practice, u can vary only slightly over distances comparable to λ so that the first two terms in the series provide an adequate approximation. Inserting the series into (13.8) gives

$$q_z = -\frac{1}{2}n\langle v\rangle\lambda\frac{\partial u}{\partial z} \tag{13.10}$$

Using the chain rule we can write

$$\frac{\partial u}{\partial z} = \left(\frac{\partial u}{\partial T}\right)\left(\frac{\partial T}{\partial z}\right)$$

Inserting this into (13.10) gives

$$q_z = -\frac{1}{2}n\langle v\rangle\lambda\left(\frac{\partial u}{\partial T}\right)\left(\frac{\partial T}{\partial z}\right) \tag{13.10a}$$

Comparing this with (13.6) shows that the thermal conductivity is given by

$$\kappa = \frac{1}{2}n\langle v\rangle\lambda\frac{\partial u}{\partial T} \tag{13.11}$$

This may be expressed in terms of more familiar quantities. Let U denote the internal energy per mole. The internal energy per atom is then $u = U/N_0$. Thus

$$\frac{\partial u}{\partial T} = \frac{1}{N_0}\frac{\partial U}{\partial T}$$

The derivative $\partial U/\partial T$ is identified as C_v, the molar heat capacity,

$$C_v = \left(\frac{\partial U}{\partial T}\right)_v$$

and thus

$$\frac{\partial u}{\partial T} = \frac{C_v}{N_0}$$

Using this result in (13.11) gives

$$\kappa = \frac{1}{2} n \langle v \rangle \lambda \frac{C_v}{N_0} \tag{13.12}$$

The factor of $\frac{1}{2}$ is not to be relied upon because of our crude average atom approach. More refined treatments lead to slightly different numerical factors. However, the underlying idea is sound; what one observes macroscopically as a heat flux is a consequence of an internal energy flux on the microscopic level.

Example 13.3 Thermal Conductivity of Air

Using the results obtained in previous chapters we can estimate the thermal conductivity of air. Under standard conditions we have

$$n \simeq 2.5 \times 10^{19} \quad \frac{\text{molecules}}{\text{cm}^3}, \qquad \langle v \rangle \simeq 3 \times 10^4 \quad \frac{\text{cm}}{\text{sec}}, \qquad \lambda \simeq 4 \times 10^{-5} \text{ cm}$$

We take $C_v = (5/2)\,R$, the value appropriate for an ideal diatomic gas. Noting that $R/N_0 = k$, the Boltzmann constant, we have

$$\frac{C_v}{N_0} = \frac{5}{2} k \simeq 3 \times 10^{-16} \quad \frac{\text{erg}}{\text{K}}$$

Inserting these values into (13.12) gives

$$\kappa \simeq 5 \times 10^3 \quad \frac{\text{erg}}{\text{K-cm-sec}} \simeq 10^{-4} \quad \frac{\text{cal}}{\text{K-cm-sec}}$$

The experimental value of κ for air at 0 °C is 0.57×10^{-4} cal/K-cm-sec, within a factor of two of our crude estimate. ✧

....

13.4.2 Compare the values of κ for air and glass. A layer of air 1-cm-thick supports a temperature difference of 20 °C. Estimate the thickness of glass needed to maintain the same temperature difference (for equal heat fluxes).

13.4.3 Assume that the atoms or molecules of a gas possess cross-sectional areas of size σ. Further, treat the gas as ideal, with f degrees of freedom. Assume a Maxwellian distribution of velocities for the purpose of determining $\langle v \rangle$. Show that

$$\kappa = \frac{fk}{4\sigma} \sqrt{\frac{8RT}{\pi M}}, \qquad M = \text{molecular weight}$$

....

REFERENCES

A kinetic theory text, selected at random, is virtually certain to contain a semiquantitative development of transport phenomena along the lines of this chapter. In particular the following texts provide such discussion:

W. Kauzmann, "The Kinetic Theory of Gases," Chapter 5. Benjamin, New York (1966).

F. L. Friedman and L. Sartori, "The Classical Atom," Chapter 2. Addison-Wesley, Reading, Massachusetts (1965).

E. H. Kennard, "Kinetic Theory of Gases," Chapter 4. McGraw-Hill, New York (1938).

Chapter 14

Mathematics for Classical Statistics

In Chapter 10 we introduced several mathematical ideas and techniques of special utility in statistical physics. The material developed there satisfied the demands of kinetic theory. Statistical mechanics makes further demands and we must now add to our mathematical repertoire.

Statistical mechanics concerns itself with the likes of atoms, molecules, and subatomic particles. Fortunately there is a common thread of statistical ideas. This is especially true in classical statistical mechanics where atoms and billiard balls are subject to the same statistical laws. The developments of this chapter are designed to guide us through classical statistical mechanics. Quantum mechanics introduces its own very special constraints and modifies the classical statistics.[†]

14.1 PROBABILITY VERSUS NUMBER OF WAYS

Figure 14.1 shows the geometrical representation of the probabilities for the possible outcomes in the game of craps. With a pair of dice there are 36 possible outcomes, each equiprobable, so that each square has an area of 1/36. From the figure it is evident how one determines the probability of rolling a four, a

[†] Classical statistics is an abbreviated appellation for classical statistical mechanics.

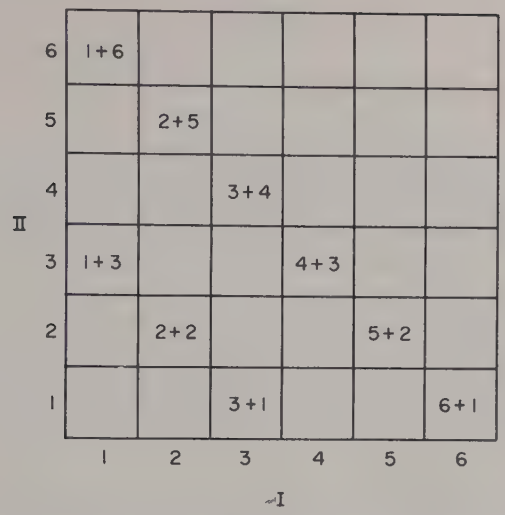


Figure 14.1. Area measures probability. The area of each square is 1/36. From the figure it is evident that three squares correspond to rolling a four, while six squares correspond to rolling a seven. The probability of rolling a four is thus $3(1/36) = 1/12$; the probability of rolling a seven is 1/6.

seven, a twelve, and so on. Precisely the same information is conveyed by letting the 36 squares be of unit area. Each square then corresponds to one possible outcome of a roll of the dice. From this viewpoint there are six ways of rolling a seven, three ways of rolling a four, and so on. In many instances an analysis in terms of the number of ways of achieving each outcome is preferred to one phrased in terms of the probability for each outcome.

....

14.1.1 The possible outcomes for the flip of a quarter are a head H or a tail T. (If the coin stands on edge, it is flipped again!) What outcomes can result from the flip of three quarters? What is the probability that flips of the three quarters result in two heads and one tail? What is the number of ways for such a result to occur?

14.1.2 An urn contains three black balls and two white balls. What is the probability that two draws will result in the withdrawal of two black balls? (One ball is withdrawn on each draw.) The first ball drawn is not returned to the urn. Enumerate all possible outcomes of the two draws. What fraction of the possible outcomes result in the withdrawal of two black balls?

....

14.2 FACTORIALS

The product of the N integers from 1 through N is called "N factorial" and written $N!$

$$N! = (1)\cdot(2)\cdot(3)\cdot(4)\cdots(N-2)\cdot(N-1)\cdot(N) \tag{14.1}$$

A more general definition of $N!$ is furnished by the relation

$$N! = \int_0^\infty x^N e^{-x}\,dx \tag{14.2}$$

The equivalence of (14.1) and (14.2) when N is a positive integer may be demonstrated by repeated integration by parts of (14.2). The garden variety factorials we ordinarily deal with are special cases of what are termed gamma functions. The gamma function $\Gamma(N)$ is defined by an integral

$$\Gamma(N) = \int_0^\infty x^{N-1} e^{-x}\,dx, \qquad N \neq 0, -1, -2, -3, \ldots \tag{14.3}$$

Comparing (14.2) and (14.3) we see that

$$N! = \Gamma(N+1) \tag{14.4}$$

For the present we need only the ordinary factorials.

....

14.2.1 Use (14.2) to show that $0! = 1$.

14.2.2 Prove that

$$N\,\Gamma(N) = \Gamma(N+1)$$

[*Hint:* Integrate by parts. Assume N to be such that all integrals converge.]

....

14.3 DISTINGUISHABLE OBJECTS AND DISTRIBUTIONS

The objects of a statistical analysis may be distinguishable in some respects but not others. A dozen white spheres may all have different radii. They are thus distinguishable on the basis of size but not color. The most direct method for sorting distinguishable objects is to number them. In the simplest situations it is possible to forget the objects and manipulate the numbers which they carry. Thus our objects are generally called 1, 2, 3, ..., and so on.

A pair of distinguishable objects, 1 and 2, allows two distinct arrangements or orderings, namely 1 2 and 2 1. For three objects 1, 2, and 3, there are six

distinct arrangements, namely

$$\begin{array}{lll} 123 & 213 & 312 \\ 132 & 231 & 321 \end{array}$$

In general there are $N!$ distinct arrangements of N distinguishable objects. To establish this fact note that the first object in line may be chosen in N ways—any one of the N objects may stand first. For each of these N choices any one of the remaining $N-1$ objects may stand second in line, so that there are a total of $N(N-1)$ distinct arrangements of the first two objects. Each of the $N(N-1)$ arrangements of the first two objects may be combined with any one of the $N-2$ possible choices for the third object for a total of $N(N-1)(N-2)$ distinct arrangements of the first three objects. The enumeration continues in this same fashion, the total number of distinct arrangements of N objects being

$$N\cdot(N-1)\cdot(N-2)\cdot(N-3)\cdots(3)\cdot(2)\cdot(1) = N!$$

Having ascertained that there are $N!$ distinct arrangements of N distinguishable objects we turn to a related problem. Suppose the N distinguishable objects are arranged in M groups. The groups themselves are also distinguishable, that is, they are numbered. For example, the N objects might be the guests in a hotel with M floors. A *distribution* of the N objects is defined by specifying how many reside in each of the M groups.

Distribution: A distribution of N objects in M groups is specified by the numbers

$$N_1, N_2, N_3, \ldots, N_M$$

where N_i is the number of objects in the ith group.

If we regard the N objects as the guests in a hotel having M floors, a distribution specifies the numbers $N_1, N_2, \ldots, N_M$ of guests on each floor. We use the symbolisms $\{N_i\}$ and $\{N_1, N_2, \ldots, N_M\}$ to designate distributions. Note that a distribution involves a certain degree of ignorance. It specifies *how many* but *not which* objects are in each group.

The *total* number of ways of achieving a distribution of N objects is $N!$. Not all of these are *distinct* ways because many of the $N!$ ways are merely rearrangements of the order *within* groups. For the distribution $\{N_1, N_2, \ldots, N_M\}$ there are a total of

$$(N_1!)(N_2!)(N_3!)\ldots,(N_M!)$$

ways of ordering objects within groups. It follows that the number of distinct ways of achieving the distribution $\{N_i\}$ is

$$D\{N_i\} = \frac{\text{total number of ways}}{\text{number of reorderings within groups}} = \frac{N!}{(N_1!)(N_2!)\cdots(N_M!)} \quad (14.5)$$

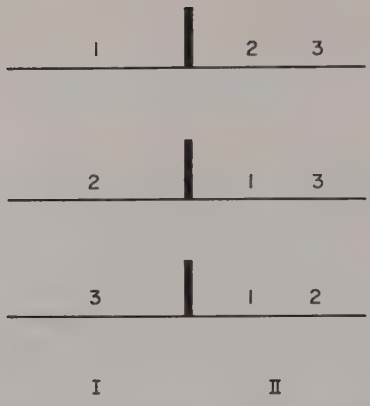


Figure 14.2. The three distinct floor assignments for the distribution; $\{N_1 = 1; N_2 = 2\}$.

To illustrate this result, consider the case $N = 3$, $M = 2$, and the distribution $\{N_1 = 1; N_2 = 2\}$. From (14.5), $D\{N_i\} = 3!/(1!)(2!) = 3$. The three distinct ways of achieving the distribution are depicted in Figure 14.2. In this instance, half of the total of $N! = 6$ arrangements are distinct—the other half correspond to rearrangements of the order within the second group. In terms of our animation, the order in which guests 2 and 3 sign the register has no bearing on the distribution if both are assigned to the same floor. Consequently $\underline{1 \blacksquare 2\,3}$ and $\underline{1 \blacksquare 3\,2}$ represent the same way of achieving the distribution $\{N_1 = 1; N_2 = 2\}$.

•••••• ••••••

14.3.1 Explain why it is the product and not the sum of the number of rearrangements within groups which gives the total number of rearrangements.

14.3.2 Consider the case $N = 4$, $M = 2$, and the distribution

$$\{N_1 = 2; N_2 = 2\}$$

Use a diagram to indicate the six distinct ways of achieving the distribution. For one of the six ways indicate the four rearrangements which give the same way of achieiving the distribution.

•••••• ••••••

Products like the one in the denominator of (14.5) occur so frequently in statistical considerations that a special notation is introduced,

$$N_1!\,N_2!\,N_3!\cdots N_M! \equiv \prod_{i=1}^{M} N_i! \tag{14.6}$$

the symbol $\prod_{i=1}^{M}$ being an instruction to form the product of M factors. In terms of the product notation,

$$D\{N_i\} = \frac{N!}{\prod_{i=1}^{M} N_i!} \tag{14.5a}$$

....

14.3.3 Evaluate the following products:

(a) $\prod_{i=1}^{4} i^2$, (c) $\prod_{i=0}^{3} 2$,

(b) $\prod_{s=1}^{3} (2s+1)$, (d) $\prod_{s=1}^{4} s^2$.

....

Next, we must generalize our development to allow for structure within the groups which comprise a distribution. For example, suppose that we refine the hotel problem so as to account for room assignments as well as floor assignments. We number the rooms on each floor and distinguish an arrangement which has, say, guest 1 in room 1 on floor 1 from one which has guest 1 in room 2 on floor 1, and so on. In a more general context we ask, "What is the total number of ways of achieving a distribution if the groups are partitioned into cells and we count as distinct arrangements which give different cell populations?" Let Q_i denote the number of cells in the ith group. If we think in terms of the hotel problem, Q_i is the number of rooms on the ith floor. The first of the N_i guests assigned to the ith floor can occupy any one of the Q_i rooms. For each—repeat—for each of these Q_i possible dispositions of the first guest, an equal number are open to the second guest, so that Q_i^2 assignments are possible for the pair.† Each time another guest is considered the existing number of assignments possible is increased by a factor of Q_i. Thus there are $Q_i, Q_i^2, Q_i^3, \ldots$, and finally, $Q_i^{N_i}$ possible room assignments for the N_i guests in the Q_i rooms on the ith floor. For the entire hotel—M floors in all—there are

$$Q_1^{N_1} Q_2^{N_2} Q_3^{N_3} \cdots Q_M^{N_M} = \prod_{i=1}^{M} Q_i^{N_i} \equiv C\{N_i\} \tag{14.7}$$

different room assignments for a given floor assignment.

....

14.3.4 Consider the case $N = 2$, $M = 2$, $Q_1 = Q_2 = 2$, and the distribution $N_1 = N_2 = 1$.

---

† Thus Q_i of these Q_i^2 assignments would place the two guests in the same room.

(a) Calculate and identify $D\{N_i\}$.

(b) For one of the distinct floor assignments identify the different room assignments.

....

This same product applies in general—$C\{N_i\}$ equals the number of different cell assignments for a given distribution.

Let $W\{N_i\}$ denote the total number of ways of achieving the distribution including the added distinctions brought in by counting different cell assignments. Evidently $W\{N_i\}$ is the product of $D\{N_i\}$ and $C\{N_i\}$. That is,

$$W\{N_i\} = \begin{pmatrix}\text{number of distinct ways of realizing a given distribution} \\ = D\{N_i\}\end{pmatrix} \cdot \begin{pmatrix}\text{number of distinct ways of distributing objects over cells} \\ \text{for the given distribution} = C\{N_i\}\end{pmatrix}$$

Combining (14.5a) and (14.7) we obtain

$$W\{N_i\} = N! \prod_{i=1}^{M} \frac{Q_i^{N_i}}{N_i!} \tag{14.8}$$

It is fair to ask just how the foregoing considerations find application in statistical physics. In Chapter 15 we consider distributions of atoms. The atoms are analogous to the hotel guests. The atoms are distributed with respect to energy. Many different energies are open to the individual atoms. The different energies correspond to the different hotel floors. It is possible for atoms to exist in different states (cells) which have equal energies. That is, the states are distinguishable on the basis of properties other than energy. The different states of equal energy correspond to different rooms on the same floor.

Statistical mechanics is concerned almost exclusively with the *most probable distribution*, which is defined as the distribution which can be formed in the largest number of ways. Thus we state as a definition:

The most probable distribution is the one for which $W\{N_i\}$ is largest.

In statistical mechanics we shall be concerned with determining the most probable distribution of particles. For example, our first venture will be to deduce the most probable spatial distribution of gas atoms. Later we will seek the most probable distribution with respect to particle energy. In all cases the problem is one of determining the most probable distribution subject to the constraints imposed by the surroundings. Generally these constraints define the thermodynamic nature of the system. For example, to represent an isolated thermodynamic system we require that the total energy and the total number

of particles have fixed values. Before pushing ahead with the physics we concern ourselves with one final mathematical topic, the saddle-point integration technique.

14.4 SADDLE-POINT INTEGRATION

In developing the formalism of statistical mechanics one frequently encounters expressions which involve the logarithms of very large factorials. Certain aspects of the theoretical analysis are greatly simplified by using Stirling's formula for such logarithms. Stirling's formula—an approximation—is

$$\ln(N!) \simeq N \ln N - N + \ln\sqrt{2\pi N}$$

Stirling's formula may be derived by a so-called saddle-point integration. The saddle-point method is quite versatile, and a worthy addition to any scientist's bag of tricks. Accordingly its development is not relegated to the appendixes.

The saddle-point method is applicable to an important class of definite integrals,

$$A = \int_a^b f(x)\, dx \tag{14.9}$$

in which the integrand $f(x)$ exhibits a sharply peaked behavior, as suggested by Figure 14.3, in which the hatched area represents the integral A. The saddle-point method approximates this area by the area under a Gaussian curve.

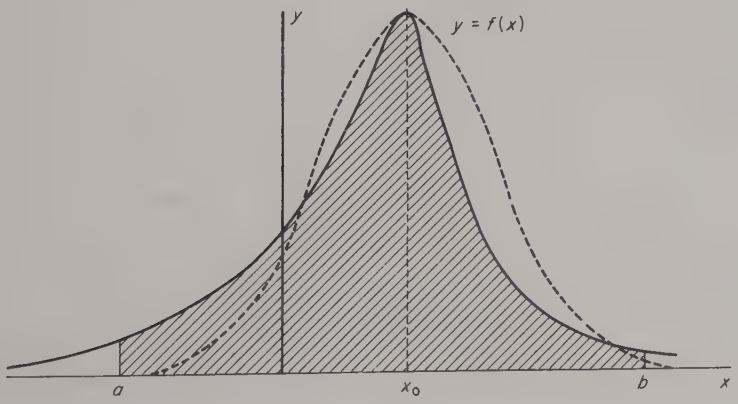


Figure 14.3. Typical form of integrand for which the saddle-point method is useful. The dashed line shows the Gaussian curve which is fitted to $f(x)$.

We begin by expressing $f(x)$ in exponential form. Let

$$g(x) = -\ln f(x) \tag{14.10}$$

so that

$$f(x) = e^{-g(x)} \tag{14.11}$$

Next, $g(x)$ is expanded in a Taylor series about the point $x = x_0$

$$g(x) = g_0 + g_0'(x-x_0) + \tfrac{1}{2}g_0''(x-x_0)^2 + \cdots \tag{14.12}$$

where

$$g_0 = g(x_0), \qquad g_0' = \left(\frac{dg}{dx}\right)_{x=x_0}, \qquad g_0'' = \left(\frac{d^2g}{dx^2}\right)_{x=x_0} \tag{14.13}$$

The first part of the approximation consists of retaining only the first three terms of the Taylor series. The success of the method depends on the precision with which the three-term series represents $g(x)$ in the range near $x = x_0$. If $f(x)$ is sharply peaked and x_0 is taken to be the peak point, the resulting approximation of $g(x)$ is generally adequate. The point x_0 is defined by the condition that $g(x)$ have a minimum there; in other words, x_0 is found by solving

$$\left(\frac{dg}{dx}\right)_{x=x_0} = 0 \tag{14.14}$$

The point x_0 is called the *saddle point* for reasons not at all in evidence here. (The method we use here is a simplification of the full saddle-point technique used in a complex variable theory.) From the relation $f(x) = e^{-g(x)}$ it is clear that $f(x)$ is a maximum where $g(x)$ is a minimum, and that the peak of $f(x)$ falls at $x = x_0$. With $g_0' = 0$, (14.12) becomes

$$g(x) = g_0 + \tfrac{1}{2}g_0''(x-x_0)^2 \tag{14.15}$$

so that

$$A = \int_a^b \exp[-g(x)]\,dx = \exp(-g_0)\int_a^b \exp[-\tfrac{1}{2}g_0''(x-x_0)^2]\,dx \tag{14.16}$$

Noting that

$$\exp[-g(x_0)] = f(x_0)$$

we can write

$$A = f(x_0)\int_a^b \exp[-\tfrac{1}{2}g_0''(x-x_0)^2]\,dx \tag{14.17}$$

The second and final approximation consists of replacing the upper and lower limits by $+\infty$ and $-\infty$. This introduces little error provided most of the area—the major contribution to the integral—lies under the central peak of the Gaussian curve. Making this change of limits and introducing the change of variable

$$z = x - x_0, \qquad dz = dx$$

transforms (14.17) into

$$A = f(x_0) \int_{-\infty}^{+\infty} \exp(-\tfrac{1}{2} g_0'' z^2)\, dz \tag{14.18}$$

The integral is of the standard Gaussian form (see Appendix A)

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} g_0'' z^2\right) dz = 2G_0\left(\frac{1}{2} g_0''\right) = \sqrt{\frac{2\pi}{g_0''}} \tag{14.19}$$

Thus the saddle-point approximation assumes the concise form

$$A = \sqrt{\frac{2\pi}{g_0''}}\, f(x_0) \tag{14.20}$$

where

$$g(x) = -\ln f(x) \tag{14.10}$$

and the saddle point x_0 is defined by the condition

$$\left(\frac{dg}{dx}\right)_{x=x_0} = 0 \tag{14.14}$$

To illustrate the method we use it to approximate two integrals which may be evaluated exactly. These examples serve to encourage a cautious optimism in the use of the saddle-point technique. As a third example we use the technique to derive Stirling's approximate formula for $\ln(N!)$.

Example 14.1

Consider the integral

$$A = \int_{-\infty}^{+\infty} \frac{dx}{x^2 + a^2}$$

The exact value of A is π/a. Here,

$$f(x) = \frac{1}{x^2 + a^2}$$

For $g(x)$ we have

$$g(x) = -\ln f(x) = +\ln(x^2+a^2)$$

and

$$g' = \frac{2x}{x^2+a^2}, \qquad g'' = \frac{2}{(x^2+a^2)} - \frac{4x^2}{(x^2+a^2)^2}$$

The condition fixing x_0 is

$$g_0' = 0 = \frac{2x_0}{x_0^2+a^2}$$

which has the solutions

$$x_0 = 0 \qquad \text{and} \qquad x_0 = \pm\infty$$

It is clear that $x_0 = 0$ is the saddle point; the solutions at $\pm\infty$ give the minima of $f(x)$. (It is always advisable to sketch the integrand roughly to aid in locating the saddle point.) Thus

$$x_0 = 0 \qquad \text{and} \qquad f(x_0) = \frac{1}{a^2}$$

Further, $g_0'' = 2/a^2$, giving for the saddle point approximation

$$A = \sqrt{\frac{2\pi}{g_0''}}\, f(x_0) = \frac{\sqrt{\pi}}{a}$$

The saddle-point approximation is too small by a factor of $\sqrt{\pi}$, the exact result being π/a. The poorest example was presented first—we hope to encourage *cautious* optimism. The saddle-point method works best when there is an honest exponential present in the integrand—an exponential which gives the integrand a sharp peak. ✧

Example 14.2

Consider next the integral

$$A = \int_0^\infty x^2 e^{-\beta x^2}\, dx$$

The exact value of this integral is $(\sqrt{\pi}/4)\,\beta^{-3/2}$ (see Appendix A). The saddle-point method proceeds from

$$f(x) = x^2 e^{-\beta x}$$

$$g(x) = -\ln f(x) = \beta x^2 - 2\ln x$$

$$g' = 2\beta x - \frac{2}{x}$$

The saddle point is determined by the equation

$$g_0' = 0 = 2\beta x_0 - \frac{2}{x_0}$$

which gives

$$x_0 = \beta^{-1/2} \qquad \text{and} \qquad f(x_0) = \beta^{-1}e^{-1}$$

Further, one finds $g_0'' = 4\beta$ and then

$$A = \frac{e^{-1}}{\sqrt{2}}\sqrt{\pi}\beta^{-3/2}$$

The numerical factor $e^{-1}/\sqrt{2}$ has the value 0.259. Thus the saddle-point method gives

$$A = 0.259\sqrt{\pi}\beta^{-3/2}$$

which compares quite favorably with the exact value

$$A = 0.250\sqrt{\pi}\beta^{-3/2}$$

This example was chosen to encourage cautious *optimism*. We now turn to a third example, a derivation of Stirling's approximation for $\ln(N!)$. ✧

Example 14.3

The Stirling approximation for $\ln(N!)$ is

$$\ln(N!) = N\ln N - N + \ln\sqrt{2\pi N} \tag{14.21}$$

To bring to bear our saddle-point machinery requires the integral representation of $N!$

$$N! = \int_0^\infty x^N e^{-x}\,dx \tag{14.2}$$

We now evaluate the integral in (14.2) via the saddle-point method. This yields the Stirling approximation. We have

$$f(x) = x^N e^{-x}$$

$$g(x) = -\ln f(x) = x - N\ln x$$

Differentiating gives

$$g' = 1 - \frac{N}{x}$$

The saddle-point equation $g_0' = 0$ gives

$$x_0 = N \qquad \text{and} \qquad f(x_0) = N^N e^{-N}$$

Further, $g_0'' = 1/N$, whereupon we find for the saddle-point approximation,

$$N! = \sqrt{\frac{2\pi}{\frac{1}{N}}} N^N e^{-N} = \sqrt{2\pi N}\, N^N e^{-N} \tag{14.22}$$

Taking logs gives the desired result

$$\ln(N!) = N \ln N - N + \ln\sqrt{2\pi N} \tag{14.21}$$

The very large numbers dealt with in statistical mechanics often justify a still cruder version of (14.21), obtained by dropping the $\ln\sqrt{2\pi N}$ term. The result, which we refer to as the *super* Stirling approximation, is

$$\ln(N!) \simeq N \ln N - N \quad \diamondsuit \tag{14.23}$$

....

14.4.1 Check the precision of the method by comparing the Stirling and super Stirling approximations with the exact value for $N = 1$, 2, and 10. Show that the percent of error decreases with increasing N by evaluating the ratios $(N!)_{\text{Stir}}/(N!)_{\text{exact}}$ and $(N!)_{\text{S.Stir}}/(N!)_{\text{exact}}$ for $N = 1$, 2, and 10.

14.4.2 Use the saddle-point method to evaluate

$$\int_0^\infty \exp\left[-\left(x^2 + \frac{a^2}{x^2}\right)\right] dx$$

The exact value is $\sqrt{\pi} e^{-2a}/2$.

14.4.3 Use the saddle-point method to show that

$$\int_0^\infty x^3 \exp[-(2x^2 + x^4)]\, dx \simeq \frac{\sqrt{\pi}}{8} \exp\left(-\frac{5}{4}\right)$$

....

Example 14.4

It often happens that the equation which determines the saddle point is transcendental, that is, it is not an algebraic equation. To illustrate the methods which may be used in this case we evaluate

$$I = \int_0^\infty \frac{x^3\, dx}{e^x - 1}$$

an integral which arises in the theory of blackbody radiation. The saddle-point

approximation for I is

$$I_{sp} = \sqrt{\frac{2\pi}{g_0''}} \frac{{x_0}^3}{e^{x_0}-1}$$

where

$$g(x) = -\ln\left(\frac{x^3}{e^x-1}\right) = -3\ln x + \ln(e^x-1)$$

The saddle-point equation is

$$g_0' = -\frac{3}{x_0} + \frac{e^{x_0}}{e^{x_0}-1} = 0$$

This may be rewritten as

$$x_0 = 3[1-e^{-x_0}]$$

This is a transcendental equation. By consulting tables of e^{-x} and exercising patience one can establish that $x_0 \simeq 2.82$. An iteration technique quickly leads to the same result. Inspection of the above equation suggests that x_0 is close to 3, that is, $1-\exp(-x_0)$ is nearly 1 when x_0 is close to 3 ($e^{-3} \simeq 0.05$). If we set $x_0 = 3$ on the right side, the saddle-point equation gives a better approximation for x_0

$$x_0 = 3(1-e^{-3}) = 2.85$$

This value of x_0 is iterated—reinserted into the exponential. This gives a still better approximation for x_0

$$x_0 = 3(1-e^{-2.85}) = 2.83$$

One more iteration gives the quoted value

$$x_0 = 3(1-e^{-2.83}) = 2.82$$

Note how quickly the iterations converge. Extreme precision in locating the saddle point is seldom warranted since the saddle-point integral is itself an approximation. In this instance the saddle-point value is 6.38, within 2% of the exact value, $\pi^4/15 \simeq 6.49$. ✧

....

14.4.4 The number of photons per unit volume (photon number density) in blackbody radiation is given by

$$\frac{N}{V} = \frac{1}{\pi^2}\left(\frac{kT}{\hbar c}\right)^3 \int_0^\infty \frac{x^2\,dx}{e^x-1}$$

The integral may be expressed in terms of the Riemann zeta function

$$\int_0^\infty \frac{x^2\,dx}{e^x-1} = 2\zeta(3) = 2.404$$

Evaluate the integral approximately using the saddle-point method. The saddle-point equation is transcendental. Use an iteration scheme to evaluate x_0.

....

REFERENCES

Brief discussions of combinatorial problems ("number of ways") will be found in:

H. Margenau and G. M. Murphy, "The Mathematics of Physics and Chemistry," 2nd ed., Chapter 12. Van Nostrand-Reinhold, Princeton, New Jersey (1961).

and

B. V. Gnedenko and A. Y. Khinchin, "An Elementary Introduction to the Theory of Probability," Chapter 5. Dover, New York (1961).

By far the "quickest" derivation of the super Stirling approximation takes advantage of the identity $\ln(N!)-\ln(N-1)! = \ln N$. See:

P. A. H. Wyatt, Elementary Statistical Mechanics without Stirling's Approximation, *J. Chem. Educ.* **39**, 27 (1962).

Chapter 15

Classical Statistical Mechanics

15.1 THE RATIONALE OF STATISTICAL MECHANICS

People who wager on the outcomes of horse races answer to many names—bettors, losers, and so on. With varying degrees of success these individuals aspire to *predict* the outcome of each race. The principles which guide the deliberations of the more proficient bettors are remindful of the rationale of statistical mechanics.

The bettor collects facts regarding each horse and its jockey, the condition of the track, and so on. Every relevant scrap of information is considered. The information is always incomplete, that is, inadequate to allow a unique determination of the outcome of a race. On the basis of incomplete information he predicts the most likely or most probable outcome of a race.

The pronouncements of statistical mechanics are also of a probabilistic nature. Statistical mechanics is based on a mathematical scheme which takes account of all available data, such as the total number of particles in a system, its volume, the temperature of its surroundings, and so on. These data are always incomplete—they constrain, but do not uniquely fix the microscopic structure of the system. Statistical mechanics singles out the most probable configuration of the system, taking care to see that its choice is consistent with the constraints.

The bettor relies on personal experience as well as the assembled facts. Statistical mechanics relies on the experience of nature. Its rules are plausible

and its predictions in accord with experiment. Of great significance is the fact that statistical mechanics permits a derivation of the laws of thermodynamics.

Briefly, the rules of statistical mechanics—to be illustrated and amplified in the following sections—are as follows:

I. Consider all distinguishable microscopic configurations of the system consistent with the constraints. In general, many different microscopic configurations result in the same *distribution.*

II. Identify the distribution corresponding to the largest number of microscopic configurations. Name this distribution *most probable.*

III. When making predictions as to the behavior of the system assume that the most probable distribution prevails. That is, the pressure, temperature, and other observable properties are assumed to be those corresponding to the most probable distribution.

We now show how adherence to these rules leads to a not unexpected prediction: The most probable configuration for a gas free of external forces is one of uniform spatial density. Consider our favorite system—a gas of N atoms. We mentally partition the container into C rooms or cells of equal volume. The cells are labeled $1, 2, \ldots, C$ according to their locations. The number of atoms occupying the ith cell is denoted by N_i. The only information we have regarding the gas is that it is composed of a definite number of atoms. There are N atoms and we know the value of N, but this is all we know about the gas. In particular, we proclaim our complete ignorance as to what the energy of the gas might be. A further implication of our limited knowledge is the absence of any net external force on the gas. A net external force would define a preferred direction in space and perhaps preferred positions in the container.

Our task is to determine the most probable distribution of the atoms. In the present context a distribution is defined by specifying the value of N_i for all C cells. According to our definition of most probable this means that we seek the distribution which can be formed in the largest number of ways. The search is subject to just one constraint, namely, that all candidates for the most probable distribution be comprised of exactly N atoms.

What we hope to find is that the most probable distribution is the one for which each cell contains the same number of atoms. Why? Because such a distribution makes the number density uniform throughout the container volume and a uniform density is one of the requirements for thermodynamic equilibrium. With N atoms in each of C cells we would have N/C atoms in each cell.

$$N_i = \frac{N}{C}, \qquad i = 1, 2, \ldots, C \tag{15.1}$$

The quantity to be maximized is the number of ways of partitioning N objects

into C groups with N_i in the ith group. This number is

$$W\{N_i\} = \frac{N!}{N_1!\,N_2!\cdots N_C!} = \frac{N!}{\prod_{i=1}^{C} N_i!} \tag{15.2}$$

This can be argued quite readily, or extracted from (14.8) by setting $Q_i = 1$ (each floor becomes 1 cell), and with C in place of the index M. The value of W for the uniform density distribution is

$$\frac{N!}{\left[\left(\frac{N}{C}\right)!\right]^C} \equiv W_{\text{mp}} \tag{15.3}$$

Direct comparison shows that W_{mp} is larger than the value of $W\{N_i\}$ for any other distribution. For example, the distribution which has $(N/C)-1$ atoms in one cell, $(N/C)+1$ in another, and N/C in each of the remaining $C-2$ cells has a value of $W\{N_i\}$ given by

$$\frac{N!}{\left(\frac{N}{C}-1\right)!\left(\frac{N}{C}+1\right)!\left[\left(\frac{N}{C}\right)!\right]^{C-2}} \equiv W'$$

Comparing W' and W_{mp} shows

$$\frac{W'}{W_{\text{mp}}} = \frac{\frac{N}{C}}{\frac{N}{C}+1} < 1$$

Thus W' is less than W_{mp} so that the uniform distribution is more probable than an adjacent slightly nonuniform distribution.† Similarly, the distribution which has $(N/C)-2$ atoms in one cell, $(N/C)+2$ in another, and N/C in each of the other cells has a value of $W\{N_i\}$ which is smaller than W_{mp} by the factor

$$\frac{\left(\frac{N}{C}-1\right)\left(\frac{N}{C}\right)}{\left(\frac{N}{C}+1\right)\left(\frac{N}{C}+2\right)} < 1$$

....

15.1.1 Consider the case of $N = 2$ atoms distributed over $C = 2$ cells of equal volume. Enumerate all distributions and verify that the uniform density distribution is most probable.

....

† Two distributions are said to be *adjacent* when their cell populations differ only slightly.

Verifying that any particular distribution that has a value of $W\{N_i\}$ less than W_{mp} does not rigorously establish the uniform distribution as most probable. It might be just a local maximum. Some "distant" distribution might have a still larger value of $W\{N_i\}$. It is somewhat like being on a mountain top, surrounded by clouds. You are certain that no adjacent points are higher, but still loftier peaks might be hidden some distance away.

The student unconvinced by the foregoing arguments will find the following demonstration more rigorous. The derivation is not presented solely in the interest of rigor. It serves mainly to introduce the ideas and methods used in Section 15.7 where the all-important *Boltzmann distribution* is deduced. As a prelude to these proceedings we establish some idea of the sizes of N and C.

Consider 1 liter of gas under standard conditions. This represents about 1/20 of a mole or $N \simeq 3 \times 10^{22}$ atoms. The C volume elements are to be small on a macroscopic scale. An optical microscope is able to resolve lengths as small as 10^{-4} cm, so that a cell volume of 10^{-12} cm^3 would be on the borderline of visibility. Let us take the cell volumes to be 10^{-14} cm^3, thereby ensuring their submacroscopic size. The characteristic atomic volume is roughly a 1 Å cube (10^{-24} cm^3) so that the cell volumes are definitely large by microscopic standards. The number of cells is

$$C = \frac{\text{container volume}}{\text{cell volume}} = \frac{10^{+3}}{10^{-14}} = 10^{17} \text{ cells}$$

Thus $N = 3 \times 10^{22}$ and $C = 10^{17}$ are large numbers, as is the average number of atoms per cell, $N/C = 3 \times 10^5$.

We now demonstrate that $W\{N_i\}$ is a maximum for the uniform distribution. The one assumption made at the outset is that

$$N_i \gg 1, \qquad i = 1, 2, \ldots, C$$

that is, we assume large cell populations. This appears quite reasonable because both N and N/C are large numbers.

When the cell populations are very large $W\{N_i\}$ changes only slightly from one distribution to an adjacent one. Consequently $W\{N_i\}$ should be essentially the same for the most probable distribution and its most adjacent neighbors. The situation is suggested by Figure 15.1 which plots $W\{N_i\}$ versus $\{N_i\}$. Since $\{N_i\}$ denotes a set of numbers, such a plot is only symbolic. Near its maximum, the interchange of one atom between two cells does not appreciably alter $W\{N_i\}$. Reasoning conversely we can use this fact to locate the maximum of $W\{N_i\}$. We write

$$W\{N_i\} = W\{N_i'\} \tag{15.4}$$

where $\{N_i\}$ is the distribution of atoms in the most probable distribution and

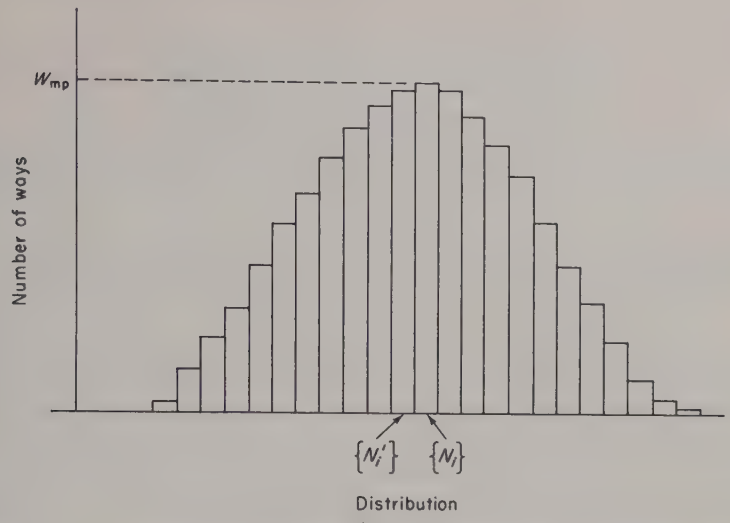


Figure 15.1. A symbolic plot suggesting the method used to determine the most probable distribution. The requirement that $W\{N_i\}$ have the same value for two adjacent distributions locates the peak.

$\{N_i'\}$ is the distribution for an adjacent distribution.† Note that $\{N_i'\}$ is immediately adjacent to $\{N_i\}$; the two distributions differ only in the population of two cells, the kth and sth. Specifically,

$$\{N_i\} = (N_1, N_2, \ldots, N_k, \ldots, N_s, \ldots, N_C) \tag{15.5}$$

and

$$\{N_i'\} = (N_1, N_2, \ldots, N_k - 1, \ldots, N_s + 1, \ldots, N_C) \tag{15.6}$$

From the way in which they have been constructed it is apparent that $\{N_i\}$ and $\{N_i'\}$ correspond to the same total number of atoms. Thus if $\{N_i\}$ is a distribution of N atoms, so also is $\{N_i'\}$

$$\sum_{i=1}^{C} N_i = \sum_{i=1}^{C} N_i' = N \tag{15.7}$$

We emphasize that k and s are arbitrary; the kth cell can be any one of the C and the sth can be any different cell. Inserting these values of N_i and N_i' into the maximization relation (15.4) gives

$$\frac{N!}{N_1!\,N_2!\cdots N_k!\cdots N_s!\cdots N_C!} = \frac{N!}{N_1!\,N_2!\cdots (N_k-1)!\cdots (N_s+1)!\cdots N_C!} \tag{15.4a}$$

† One does essentially the same thing in an ordinary maxima–minima problem. Thus, for example, requiring that $f(x) = f(x+dx)$ locates the maxima and/or minima of $f(x)$—the points at which $df/dx = 0$.

Canceling common factors and noting that

$$\frac{(N_s+1)!}{N_s!} = N_s + 1, \qquad \frac{N_k!}{(N_k-1)!} = N_k$$

reduces (15.4a) to

$$N_s + 1 = N_k \tag{15.8}$$

Because the cell populations are large we can replace N_s+1 by N_s with the result that (15.8) becomes

$$N_s = N_k \tag{15.9}$$

However, k and s are arbitrary; (15.9) tells us that for the most probable distribution the populations of *any* two cells we might choose must be the same. This means that all cells must contain the same numbers of atoms. Thus we have followed a different path to a determination that the uniform distribution is most probable. A parallel path is followed to derive the famous Boltzmann distribution law.

In the next section we introduce the important concept of phase space. This will permit us to move from the cells of ordinary coordinate space used in this section to the cells of phase space with which statistical mechanics deals.

15.2 PHASE SPACES

In Chapter 11 we introduced the notion of velocity space by analogy with coordinate space. With the abstraction of velocity space now firmly entrenched it is not difficult to visualize a combined position–velocity space. Such a space has six dimensions. A point in the space is specified by six quantities

$$x, \quad y, \quad z, \quad v_x, \quad v_y, \quad v_z$$

The combined position–velocity space is one type of phase space. Geometrically, a point in phase space fixes the phase of particle motion; where it is and how it is moving. It is not possible to plot or visualize points in six dimensions. Nevertheless, phase space provides an elegant framework for the description of particle behavior. Moreover, phase space is essential for the logical development of many segments of statistical mechanics. Phase space is both economical and essential.

The motion of a particle describes a real trajectory in coordinate space. With each point along the path of the particle we associate a position $\mathbf{r}$ and a velocity $\mathbf{v}$. For each point on the real path there corresponds a representative point in the six-dimensional phase space $(\mathbf{r}, \mathbf{v})$. As the particle moves, its representative point also changes and we can speak of a trajectory in phase space.

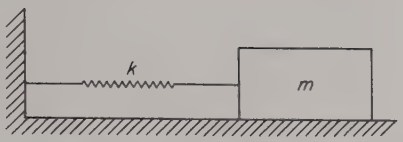


Figure 15.2. One-dimensional simple harmonic oscillator. The symbol k refers to the spring constant, not the Boltzmann constant.

The particle moves in a real three-dimensional space; its representative point moves in a six-dimensional phase space.

Phase-space trajectories can be pictured in two dimensions (on these pages!) if we deal with a system undergoing one-dimensional motion. The phase space then has two dimensions with, say, x and v_x denoting the coordinates of the representative point. As an example, consider the one-dimensional simple harmonic oscillator (SHO) of Figure 15.2. The motion of the oscillator can be described by

$$x(t) = A\cos\omega t, \qquad \omega^2 = \frac{k}{m} \tag{15.10}$$

$$v_x(t) = -\omega A\sin\omega t \tag{15.11}$$

From these equations we find

$$\frac{x^2}{A^2} + \frac{v_x^{\,2}}{A^2\omega^2} = 1 \tag{15.12}$$

This, in turn, shows that the phase-space trajectory of the representative point is an ellipse, as depicted in Figure 15.3. The arrows on the ellipse indicate the sense in which the representative point traverses the phase-space trajectory.

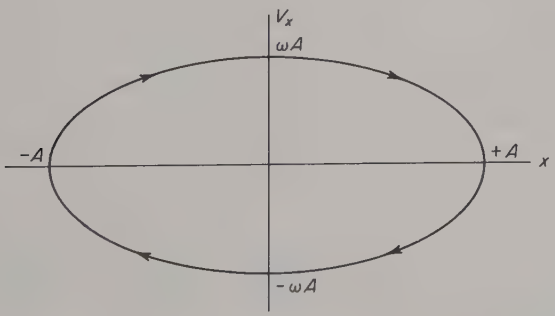


Figure 15.3. Phase-space trajectory of representative point for the simple harmonic oscillator. The representative point circulates in clockwise fashion around the ellipse.

....

15.2.1 Explain why the phase-space trajectory of Figure 15.3 is traced clockwise rather than counterclockwise. [*Hint:* Start at $t = 0$ and consider the signs of x and v_x as time elapses.]

15.2.2 A particle undergoes one-dimensional motion, being subject to a constant acceleration g. The initial velocity is v_0. Determine $x(t)$ and $v_x(t)$. Eliminate t to obtain an equation relating x and v_x. Plot the resulting phase-space trajectory for $g = \frac{1}{2}$ and $v_0 = -1$. Locate the initial point ($t = 0$) and indicate the subsequent path of the representative point along the trajectory. The initial position of the particle is $x(t = 0) = 0$.

15.2.3 A rubber ball bounces in such a fashion that it rebounds to a height of 9/10 that to which it rose on the preceding bounce. *Sketch* its phase-space trajectory, assuming that its velocity changes discontinuously at impact, that is, that it remains in contact with the wall for zero time.

....

Particles with coordinates in d^3r and velocities in d^3v have representative points which cluster closely in phase space. It is natural to introduce a phase-space volume element. We refer to the six-dimensional element

$$d^3r\, d^3v = dx\, dy\, dz\, dv_x\, dv_y\, dv_z \tag{15.13}$$

as the phase-space volume element.

For systems in thermodynamic equilibrium the density must be uniform. It is then unnecessary to deal with small spatial volume elements—the observable properties of the system are the same at all points. This permits us to deal with a large volume element—the total volume of the system. The corresponding phase-space volume element is then

$$d^3r\, d^3v = V\, dv_x\, dv_y\, dv_z \tag{15.14}$$

For reasons which must be left unstated at this point, a slightly different version of phase space is generally employed in statistical mechanics. Points in phase space are specified in terms of position and linear momentum rather than in terms of position and velocity. In many situations the difference between momentum and velocity is merely a factor of m, the particle mass. Thus

$$\mathbf{p} = m\mathbf{v}, \qquad dp_x = m\, dv_x$$

Such a simple relation between momentum and velocity does not always hold. For example, a photon traveling at speed c has zero rest mass. However, the photon momentum is not zero. The magnitude of the photon momentum follows from the relation

$$\varepsilon^2 = (pc)^2 + (mc^2)^2 \tag{15.15}$$

where ε is the energy, p is the momentum, and m is the rest mass. The rest mass of the photon is zero so that

$$\varepsilon = pc \tag{15.16}$$

In situations where we deal with the position–momentum phase space, a representative point is labeled by the six variables

$$x, \quad y, \quad z, \quad p_x, \quad p_y, \quad p_z$$

Velocity space and the accompanying volume element d^3v are replaced by momentum space and the volume element d^3p. In Cartesian momentum space

$$\mathbf{p} = (p_x, p_y, p_z), \qquad d^3p = dp_x\, dp_y\, dp_z \tag{15.17}$$

A description in terms of spherical momentum variables is often desired. The momentum can be specified by giving its magnitude p and direction (θ, φ). Volume elements needed for spherical momentum space are

$$d^3p = p^2\, dp \sin\theta\, d\theta\, d\varphi \quad \text{and} \quad d^3p = 4\pi p^2\, dp \tag{15.18}$$

The volume element in the position–momentum phase space is

$$d^3r\, d^3p = \begin{cases} \text{volume of position–momentum phase space} \\ \text{containing representative points of particles} \\ \text{with positions in } d^3r \text{ and momenta in } d^3p \end{cases}$$

Having identified phase space as the proper setting for statistical mechanics we proceed to develop the theory in some detail.

15.3 PHASE SPACES AND HOTELS: AN ANALOGY

In Chapter 14 we introduced some of the mathematical ideas needed to develop statistical mechanics. The treatment there was animated in terms of hotel guests, room and floor assignments, and so on. Our attention shifts from hotel guests to atoms and molecules. Certain ideas are most clearly presented with reference to a specific physical system. For this role we choose our favorite of long standing, a gas of N atoms confined within a volume V. The gas acts only as a familiar frame of reference; the results are of wide applicability within the framework of classical statistical mechanics. Our analysis is restricted to systems in thermodynamic equilibrium. Thus in the case of the gas, its pressure, density, and temperature are spatially uniform and do not change with time.

The results established in the preceding chapter enable us to form helpful correspondences between hotels and their guests and phase space and its occupants. The atoms are analogous to the hotel guests, while the complete phase space corresponds to the hotel. To construct the analogs of hotel rooms

we mentally partition the phase space into a great many cells. Each phase-space cell is a suitable home for one or more atoms. Our discussion of hotels was concerned with both room and floor designations. Some care is required to establish the phase-space counterpart of a hotel floor assignment. With each phase-space cell there is associated a definite energy—the energy of an atom in that cell.† An important assumption is implicit in assigning a definite energy to each segment of phase space. To say that an atom has a certain energy because it is in a particular region of phase space ignores the effects of other atoms. The proximity of other atoms is partially responsible for determining the total energy of any given atom. Thus our basic approach here is limited to situations in which mutual interactions can be ignored. We deal with an independent-particle model. Such a model proves adequate in a number of cases. For example, in Chapter 11 such a model led us to the ideal gas law. Statistical mechanics copes with the problem of interacting particles by introducing more sophisticated phase spaces. Happily, these more powerful approaches are conceptual extensions of the ideas we now develop in connection with the independent-particle model.

Let us imagine a line along which the phase-space cell energies are plotted. We refer to this energy-axis pattern as the energy spectrum for the system. The nature of the physical system determines the details of an energy spectrum. However, certain characteristics are common to all energy spectra. The most important feature is that large numbers of phase-space cells have nearly equal energies. We may express this idea in a more quantitative way through the relation

$$\frac{\Delta\varepsilon}{\varepsilon} \ll 1$$

where ε is a representative cell energy and $\Delta\varepsilon$ a typical energy difference between adjacent cells. The high density of cells along the energy axis is very important. When the energy spacing of cells is small by comparison with the energies themselves we may legitimately group several cells and assign each the same energy. This grouping enables us to deal with wholesale numbers of atoms having nearly equal energies. We can pursue this idea through the hotel analogy. In any large hotel there are many rooms which rent for the same price even though the characteristics are not exactly the same for each room. Some rooms renting for \$15.00 a day will be slightly larger than others, some will have newer furniture, and so on. It would be very cumbersome to have rents of \$15.05, \$14.98, \$15.12, \$15.07, ..., even though these figures might represent

† Once again we economize on words by stating that the atom is in a phase-space cell, when in fact it is the representative point of the atom.

the relative merits of the rooms. An average figure of \$15.00 is far more expedient. To complete the analogy between phase space and hotels we must group rooms of the same rent on the same floor. We play the same game in statistical physics, grouping phase-space cells with nearly equal energies and assigning each an average energy. A group of cells with the same energy corresponds to a floor of rooms which rent for the same rate.

In summary, we have the following analogies between phase space and hotels: The complete phase space corresponds to a hotel. Individual phase-space cells are analogs of hotel rooms. Rooms which command equal rent are grouped on the same floor. Phase-space cells whose energies are nearly the same are considered in groups, all cells in the same group being assigned equal energy.

15.4 MICROSTATES AND MACROSTATES

The concepts of *microstates* and *macrostates* of a physical system are of great importance in statistical mechanics. Both microstates and macrostates are defined with reference to the populations of groups of phase-space cells.

Let Q_i denote the number of phase-space cells in the ith group and let ε_i be the corresponding cell energy. Hotels generally have some rooms occupied and others empty. Phase space is very similar. There will be both occupied and empty phase cells. Let N_i denote the total number of atoms which reside in the Q_i cells. Further, let M denote the total number of groups of cells. The energy axis is thereby partitioned into M segments, the ith segment comprising Q_i cells which contain a total of N_i atoms. A *macrostate* of the system is defined by specifying the populations of the M groups of cells. Thus a macrostate is determined by a specification of the numbers $N_1, N_2, N_3, \ldots, N_M$, that is, by the numbers of atoms in each group of cells. Evidently a *macrostate* is a *distribution* in phase space. Note in particular that a macrostate specifies only how many atoms, and not which atoms occupy each group of cells. A *microstate* is defined by specifying which phase-space cell each atom occupies. As such, a microstate corresponds to one way of realizing a particular distribution of hotel guests. Table 15.1 summarizes the analogies between hotel guests and the occupants of phase-space cells.

We have previously stressed the point that the measurable properties of the gas characterize the average behavior of many atoms. The measurable properties depend on the numbers of atoms occupying the various phase-space cells. The observable properties do not depend on which atoms these might be. A microstate indicates "which ones" as well as "how many" and thus supplies more information than is needed to determine the measurable properties of the system. A less detailed description of the gas, but one adequate to explain its

TABLE 15.1

Guest–Atom Analogies

Guest	Atom
N guests reside in a hotel with M floors, there being Q_i rooms on the ith floor. All rooms on the same floor rent at the same rate.	N atoms are located in M groups of phase-space cells, there being Q_i cells in the ith group. Each cell in a group corresponds to the same energy.
Specifying the numbers of guests $N_1, N_2, \ldots, N_M$ on each floor defines a *distribution.*	Specifying the numbers of atoms $N_1, N_2, \ldots, N_M$ in each group defines a *macrostate.*
A *way* of realizing a particular distribution is defined by identifying which room each guest occupies.	A *microstate* corresponding to a particular macrostate is defined by identifying which cell each atom occupies.

measurable properties, is provided by the macrostate. If the measurable properties depend only on the macrostate of the gas it might seem that the concept of a microstate is superfluous. However, this is not the case—we need microstates. In particular, microstates are essential to the development of the notion of thermodynamic probability.

15.5 THERMODYNAMIC PROBABILITY

Visualize some microstate of the gas. This microstate corresponds to some macrostate of the gas. In general, a great many different microstates correspond to the same macrostate. A macrostate is a distribution of the atoms, while each microstate giving rise to it constitutes a distinct way of achieving the distribution. The number of different microstates which result in the same macrostate is readily deduced. In fact, the direct analogies between atoms and hotel guests let us recognize $W\{N_i\}$ as the number of microstates giving rise to the same macrostate

$$W\{N_i\} = N! \prod_{i=1}^{M} \frac{Q_i^{N_i}}{N_i!} \tag{14.8}$$

Here the N objects are atoms, distributed over M groups of phase-space cells, such that the ith group of Q_i cells contains N_i atoms.

Microstates consistent with all of the macroscopic constraints imposed on the system by its surroundings are referred to as *accessible*. In the case of our gas, for example, microstates which have one or more atoms outside the container are conceivable but not accessible. Similarly, if we specify the total

energy of the gas to be 16 J, the accessible microstates are restricted to be those which result in this energy. Other microstates are conceivable, but not accessible under the constraints imposed.

The macroscopic constraints imposed by the surroundings of a system determine its thermodynamic state. In general, the thermodynamic state of a system admits an enormous number of different macrostates. That is, a great many different macrostates can result in the same thermodynamic state. As we have just seen, there are also numerous microstates for each macrostate, $W\{N_i\}$ of them. Figure 15.4 suggests the fragmenting association of thermodynamic state, macrostates, and microstates.

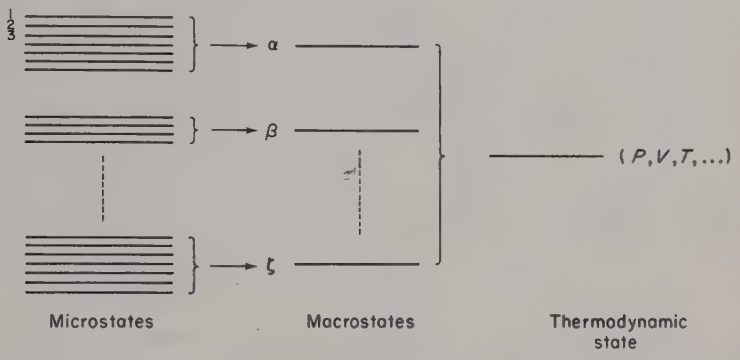


Figure 15.4. A great many accessible microstates correspond to the same macrostate. Similarly, a vast number of macrostates give rise to the same thermodynamic state.

At this point we introduce a crucial statistical hypothesis—a *postulate* of statistical mechanics.

All accessible microstates are equiprobable

The basis for advancing the postulate is quite simple; we have no reason—no evidence—which favors one accessible microstate over any other. The postulate is little more than Laplace's principle of insufficient reason. Assuming that all accessible microstates are equiprobable it follows that the probability of finding a system in a particular macrostate is directly proportional to $W\{N_i\}$, the number of microstates which result in that macrostate. In statistical mechanics $W\{N_i\}$ is generally referred to as the *thermodynamic probability* of the macrostate.

$$\text{thermodynamic probability of finding a system in the macrostate specified by the distribution } \{N_i\} \equiv W\{N_i\} = N! \prod_{i=1}^{M} \frac{Q_i^{N_i}}{N_i!} \tag{15.19}$$

Although $W\{N_i\}$ is not a probability in the usual sense, it does convey the same information as a properly normalized probability.[†]

15.6 THE MOST PROBABLE MACROSTATE AND THERMODYNAMIC EQUILIBRIUM

Knowing which macrostate a system occupies allows one to compute mean values and thus to predict values for its measurable properties. Our analysis will reveal only the most probable macrostate, the macrostate for which the thermodynamic probability is a maximum. However, the system cannot always be in the most probable macrostate. The motions of the atoms carry the system through a succession of microstates. Not all of the microstates through which the system passes can correspond to the most probable macrostate. Such a feat would mean that external (macroscopic) constraints were acting to control the motions of the atoms. It would mean that Maxwell demons are possible. The fact that no such control is possible forms the basis for assuming that all accessible microstates are equiprobable. We conclude that the macrostate of the system changes with time. On the other hand the fact that we deal with a system in thermodynamic equilibrium means that the measurable properties remain virtually constant.[‡] This pair of observations leads to an important conclusion: The measurable properties of the system are nearly the same for a vast majority of the macrostates. The plausibility of this judgement is strengthened by considering its antithesis. If the measurable properties varied markedly with changes in the macrostate, a system could never reach thermodynamic equilibrium. Changes in the macrostate occur billions upon billions of times each second. The system would be unable to settle down into an equilibrium state if the observable properties changed appreciably with changes in the macrostate. The realization that the system continually changes macrostates, coupled with the existence of thermodynamic equilibrium leads us to believe that the measurable properties of the system are essentially the same for a sizable majority of all macrostates. Not every macrostate, because we can observe fluctuations. The fluctuations arise whenever the system ventures into a series of markedly offbeat macrostates. Because the most probable macrostate is representative of virtually all macrostates, statistical mechanics directs us to assume that

---

[†] Here is another situation where specifying the number of ways is equally as informative as giving the probability.

[‡] The measurements may reveal fluctuations, the variations about equilibrium values. We can picture a fluctuation as a succession of nonequilibrium macrostates.

The most probable macrostate corresponds to the thermodynamic equilibrium state.

By corresponds we mean that the macroscopic properties—pressure, temperature, internal energy, and so on—have the same values for the thermodynamic equilibrium state and the most probable macrostate.

15.7 THE BOLTZMANN DISTRIBUTION

We now seek the most probable macrostate for an isolated system in thermodynamic equilibrium. An isolated system can exchange neither particles nor energy with its surroundings. Therefore, the total number of particles N and the internal energy U of the system are constants. Macrostates aspiring to gain the status of most probable must specify a distribution of N atoms with a total energy equal to U.

The thermodynamic probability $W\{N_i\}$ as expressed by (15.21) is a function of the M variables $N_1, N_2, N_3, \ldots, N_M$. The most probable macrostate is determined by an extension of the approach used in Section 15.2. Let $\{N_i\}$ denote the most probable macrostate[†]

$$\{N_i\} = (N_1, N_2, \ldots, N_k, \ldots, N_s, \ldots, N_M) \tag{15.20}$$

and let $\{N_i'\}$ denote the macrostate immediately adjacent to $\{N_i\}$. Specifically, $\{N_i'\}$ is to differ from $\{N_i\}$ only in the populations of two groups of cells, the kth and sth:

$$\{N_i'\} = (N_1, N_2, \ldots, N_k - 1, \ldots, N_s + 1, \ldots, N_M) \tag{15.21}$$

We emphasize that k and s are arbitrary; the kth group of cells can be any one of the M and the sth any one of the remaining groups. The adjacent macrostate $\{N_i'\}$ can evidently be formed from the most probable macrostate $\{N_i\}$ by the transfer of one atom from the kth to the sth group of cells. This guarantees that both macrostates correspond to the same number of atoms

$$\sum_{i=1}^{M} N_i = \sum_{i=1}^{M} N_i' = N \tag{15.22}$$

Let W and W' denote the thermodynamic probabilities for $\{N_i\}$ and $\{N_i'\}$, respectively. The most probable macrostate is located by requiring that it have the same thermodynamic probability as its immediately adjacent neighbors

$$W = W' \tag{15.23}$$

[†]Henceforth, $\{N_i\}$ will be referred to as a macrostate instead of a distribution.

The idea behind (15.23) is that the macrostate having the largest thermodynamic probability differs in the most minimal way from adjacent macrostates. The differences are so slight that the most probable and immediately adjacent macrostates have essentially equal thermodynamic probabilities. Reasoning conversely, we can use this equality, as expressed by (15.23), to identify the most probable macrostate.

If the N_i atoms in the ith group of cells each have energy ε_i, their combined energy is $N_i\varepsilon_i$. The total internal energy of the N atoms distributed over M groups of cells is $\sum_{i=1}^{M} N_i\varepsilon_i$. In the competing macrostate $\{N_i'\}$ there are N_i' atoms in the ith group of cells so that the total internal energy is $\sum_{i=1}^{M} N_i'\varepsilon_i$. To satisfy the constraint

$$\text{total internal energy of gas} = U$$

subjects the macrostates $\{N_i\}$ and $\{N_i'\}$ to the constraint

$$U = \sum_{i=1}^{M} N_i\varepsilon_i = \sum_{i=1}^{M} N_i'\varepsilon_i \tag{15.24}$$

The distributions of atoms producing $\{N_i\}$ and $\{N_i'\}$ are given by (15.20) and (15.24) so that both macrostates correspond to the same total number of atoms. We now have a well-posed mathematical problem: Determine the form of the N_i which simultaneously satisfy (15.23) and (15.24). If we insert the explicit forms of $W\{N_i\}$ and $W\{N_i'\}$ given by (15.19), (15.23) becomes

$$W = N!\prod_{i=1}^{M}\frac{Q_i^{N_i}}{N_i!} = N!\prod_{i=1}^{M}\frac{Q_i^{N_i'}}{N_i'!} = W' \tag{15.23a}$$

Whereas (15.23a) is a relation involving products of functions of the N_i, (15.24) is a relation between sums of the N_i. For this reason it is mathematically more convenient to replace (15.23) by

$$\ln W = \ln W' \tag{15.25}$$

This is equivalent to requiring $W = W'$ because the macrostate for which $\ln W$ is a maximum is also the macrostate for which W is largest. Furthermore, $\ln W$ involves a sum of functions of the N_i, as does U.

....

15.7.1 Show that the function

$$\varphi(x) = x^3\exp(-x^2)$$

and $\ln\varphi(x)$ attain their maximum values at the same point.

....

There is also a very strong physical reason for choosing to maximize $\ln W$ rather than W itself. As we shall discover, $\ln W$ is directly proportional to the entropy of a thermodynamic state, and thermodynamics has taught us that the state of maximum entropy is the equilibrium state. Thus by maximizing $\ln W$ we locate the state of largest entropy, the equilibrium state. In order to guarantee that both (15.24) and (15.25) are satisfied simultaneously we write

$$\ln W - \beta \sum_{i=1}^{M} N_i \varepsilon_i = \ln W' - \beta \sum_{i=1}^{M} N_i' \varepsilon_i \tag{15.26}$$

where $-\beta$ is a constant with the dimensions of (1/energy) introduced to render $-\beta U$ dimensionless. Thus both $\ln W$ and $-\beta U$ are dimensionless quantities, and (15.26) holds only when both (15.24) and (15.25) are satisfied. We will show later that β is the reciprocal of the characteristic thermal energy kT

$$\beta = \frac{1}{kT} \tag{15.27}$$

Inserting the explicit forms for W and W' into (15.26), and using the fact that

$$\ln\left(\prod_i X_i\right) = \sum_i \ln X_i \tag{15.28}$$

results in

$$\ln N! + \sum_{i=1}^{M}\left[\ln\left(\frac{Q_i^{N_i}}{N_i!}\right) - \beta N_i \varepsilon_i\right] = \ln N! + \sum_{i=1}^{M}\left[\ln\left(\frac{Q_i^{N_i'}}{N_i'!}\right) - \beta N_i' \varepsilon_i\right] \tag{15.29}$$

The most probable macrostate and the immediately adjacent macrostate differ only in the populations of the kth and sth groups of cells

$$N_k' = N_k - 1$$

$$N_s' = N_s + 1$$

For the other $M-2$ groups of cells

$$N_i' = N_i$$

Canceling identical factors in (15.29) leaves

$$\ln\left(\frac{Q_k^{N_k}}{N_k!}\right) - \beta N_k \varepsilon_k + \ln\left(\frac{Q_s^{N_s}}{N_s!}\right) - \beta N_s \varepsilon_s = \ln\left(\frac{Q_k^{N_k-1}}{(N_k-1)!}\right) - \beta(N_k-1)\varepsilon_k + \ln\left(\frac{Q_s^{N_s+1}}{(N_s+1)!}\right) - \beta(N_s+1)\varepsilon_s$$

Grouping k and s terms on opposite sides of the equation and noting that, for example,

$$\ln\left(\frac{Q_k^{N_k-1}}{(N_k-1)!}\right) - \ln\left(\frac{Q_k^{N_k}}{N_k!}\right) = \ln\left(\frac{N_k}{Q_k}\right)$$

results in

$$\ln\left(\frac{N_s+1}{Q_s}\right) + \beta\varepsilon_s = \ln\left(\frac{N_k}{Q_k}\right) + \beta\varepsilon_k$$

Assuming that the population of each group of cells is large allows us to replace N_s+1 by N_s, yielding

$$\ln\left(\frac{N_s}{Q_s}\right) + \beta\varepsilon_s = \ln\left(\frac{N_k}{Q_k}\right) + \beta\varepsilon_k \tag{15.30}$$

However, k and s are arbitrary; hence the quantity $\ln(N_i/Q_i) + \beta\varepsilon_i$ must be the same for all M groups of cells,

$$\ln\left(\frac{N_i}{Q_i}\right) + \beta\varepsilon_i = \text{constant},^{\dagger} \qquad i = 1, 2, \ldots, k, \ldots, s, \ldots, M$$

Calling the constant $\ln A$, where A is itself constant, we arrive at the distribution

$$N_i = AQ_i e^{-\beta\varepsilon_i}, \qquad \beta = \frac{1}{kT} \tag{15.31}$$

This is the famous *Boltzmann distribution*. It is the basis for a vast number of applications of statistical mechanics. The constant A may be expressed in terms of N and β by noting that the Boltzmann distribution satisfies the equation of constraint $N = \sum_{i=1}^{M} N_i$. Thus

$$N = \sum_{i=1}^{M} N_i = A \sum_{i=1}^{M} Q_i \exp(-\beta\varepsilon_i)$$

If we define

$$Z(\beta) \equiv \sum_{i=1}^{M} Q_i \exp(-\beta\varepsilon_i) \tag{15.32}$$

we have

$$A = \frac{N}{Z(\beta)}$$

---

† In general the constant will depend on the thermodynamic state of the system, its temperature, volume, and so on. It is a constant only in the sense that it has the same value for all M phase-space cells.

The populations may now be expressed as

$$N_i = \frac{NQ_i \exp\left(\frac{-\varepsilon_i}{kT}\right)}{Z(\beta)}, \qquad \beta = \frac{1}{kT} \tag{15.33}$$

Observe that N_i/N, the fraction of the atoms which have energy ε_i, is directly proportional to the exponential $\exp(-\varepsilon_i/kT)$. This all-important quantity governing the cell populations is known as the Boltzmann factor:

$$e^{-\varepsilon/kT} \Rightarrow \text{Boltzmann factor}$$

It is essential to understand how $e^{-\varepsilon/kT}$ behaves as ε and/or kT change. The Boltzmann factor reveals that there is a competition for the available internal energy. For a fixed temperature the Boltzmann factor decreases as the energy ε is increased. A modest change in the ratio ε/kT changes the Boltzmann factor dramatically. For example,

$$\varepsilon = 0, \qquad \exp\left(\frac{-\varepsilon}{kT}\right) = e^{-0} = 1$$

$$\varepsilon = kT, \qquad \exp\left(\frac{-\varepsilon}{kT}\right) = e^{-1} = 0.367$$

$$\varepsilon = 10kT, \qquad \exp\left(\frac{-\varepsilon}{kT}\right) = e^{-10} \simeq 5 \times 10^{-5}$$

$$\varepsilon = 30kT, \qquad \exp\left(\frac{-\varepsilon}{kT}\right) = e^{-30} \simeq 10^{-13}$$

The Boltzmann factor is a sharply decreasing function of energy. The majority of atoms have energies comparable to kT. Atoms with energies over $10kT$ are exceptional and constitute only a small fraction of all atoms. The quantity kT emerges as the characteristic thermal energy.

For a given energy, the Boltzmann factor decreases as T decreases. As a consequence a decrease in temperature shifts the population toward phase-space cells of lower energy.

We remark in passing that classical statistical mechanics (CSM) is often referred to as Boltzmann statistics, especially in situations where its predictions are contrasted with those of quantum statistical mechanics.

The quantity

$$Z(\beta) = \sum_{i=1}^{M} Q_i \exp(-\beta\varepsilon_i) \tag{15.32}$$

answers to several names, Zustandsumme and partition function being the most widely used. We use the latter.

Example 15.1 The Two-Level System

The so-called two-level system has only two available phase-space cells ($Q_1 = Q_2 = 1$). The energies associated with the two cells are

$$\varepsilon_1 = 0, \qquad \varepsilon_2 = +\varepsilon, \qquad \varepsilon > 0$$

With ε positive, ε_1 is the lower of the two energies. The partition function for the two level system is

$$Z(\beta) = \sum_{i=1}^{2} Q_i \exp(-\beta\varepsilon_i) = \exp(-\beta \cdot 0) + \exp(-\beta \cdot \varepsilon)$$

or

$$Z(\beta) = 1 + \exp(-\beta\varepsilon) \quad \diamondsuit$$

....

15.7.2 (a) With reference to Example 15.1 show that the fractional populations of the two cells are

$$\frac{N_1}{N} = \frac{1}{1+e^{-\beta\varepsilon}}, \qquad \frac{N_2}{N} = \frac{e^{-\beta\varepsilon}}{1+e^{-\beta\varepsilon}}$$

(b) With ε = constant = 1 eV, determine the population ratio N_2/N_1 for the three cases,

$$kT = 10 \text{ eV}, \qquad kT = 1 \text{ eV}, \qquad kT = 0.1 \text{ eV}$$

How do your results support the statement that "a decrease in temperature shifts the population toward phase-space cells of lower energy."

....

15.8 THE PARTITION FUNCTION FOR AN IDEAL GAS

The parameter β is determined implicitly by (15.24) and (15.33), from which we find

$$\frac{U}{N} = \sum_{i=1}^{M} \frac{Q_i \varepsilon_i \exp(-\beta\varepsilon_i)}{Z(\beta)} \tag{15.34}$$

The right side of (15.34) can be expressed as a derivative of $\ln Z(\beta)$. Note first that

$$-\frac{\partial}{\partial\beta} \ln Z(\beta) = -\frac{1}{Z(\beta)} \cdot \frac{\partial Z}{\partial\beta}$$

From the definition of the partition function $Z(\beta)$ we have

$$\frac{\partial Z}{\partial \beta} = \frac{\partial}{\partial \beta} \sum_{i=1}^{M} Q_i \exp(-\beta \varepsilon_i) = -\sum_{i=1}^{M} Q_i \varepsilon_i \exp(-\beta \varepsilon_i)$$

so that

$$-\frac{\partial}{\partial \beta} \ln Z(\beta) = \sum_{i=1}^{M} \frac{Q_i \varepsilon_i \exp(-\beta \varepsilon_i)}{Z(\beta)} \tag{15.35}$$

Comparing (15.34) and (15.35) shows that

$$\frac{U}{N} = -\frac{\partial}{\partial \beta} \ln Z(\beta) \tag{15.36}$$

While this diabolical maneuver simplifies the form of the equation relating β and U/N it still does not provide an explicit equation of the form

$$\beta = \text{some explicit function of } \frac{U}{N}$$

We now show that $\beta = 1/kT$. The scheme is to evaluate $Z(\beta)$ for an ideal monatomic gas, which leads to

$$\frac{U}{N} = \frac{3}{2} \frac{1}{\beta} \tag{15.37}$$

We have previously shown that the internal energy of an ideal gas is related to its temperature by

$$\frac{U}{N} = \frac{3}{2} kT \tag{15.38}$$

from which it follows that $\beta = 1/kT$.

Although the notation suggests that $Z(\beta)$ depends on β alone this is most definitely not the case. In general the partition function depends on the full complement of thermodynamic variables. However, to avoid unwieldy notations we write simply $Z(\beta)$.

To evaluate $Z(\beta)$ we must establish the connection between the number of cells Q_i and the cell energy ε_i. This is done here for an ideal monatomic gas. Our original decision to consider groups of cells was prompted by the fact that many particles have nearly equal energies. It also means that there are many phase-space cells corresponding to a small range of energy. Suppose we consider the energy range from ε to $\varepsilon + d\varepsilon$, where $d\varepsilon$ is macroscopically small, but

microscopically large. That is, there are many cells whose energies fall in the range from ε to $\varepsilon + d\varepsilon$. The high density of cells along the energy axis allows us to replace the summation over groups of cells by an integration over phase space. We make the replacement

$$Z(\beta) = \sum_{i=1}^{M} Q_i \exp(-\beta\varepsilon_i) = \int_{\substack{\text{phase} \\ \text{space}}} dQ(\varepsilon) \exp(-\beta\varepsilon) \tag{15.39}$$

where $dQ(\varepsilon)$ is the number of phase-space cells with energy in the range from ε to $\varepsilon + d\varepsilon$.

....

15.8.1 Show how the sums

$$\sum_{n=0}^{N} n = \frac{N(N-1)}{2}$$

and

$$\sum_{n=0}^{N} n^2 = \frac{N(2N^2-3N+1)}{6}$$

are closely approximated by integrals when $N \gg 1$.

....

For an ideal monatomic gas the energy of an atom is purely translational kinetic energy. The energy and speed are related by

$$\varepsilon = \tfrac{1}{2}mv^2 \tag{15.40}$$

Cells with energies in the range ε to $\varepsilon + d\varepsilon$ correspond to a particular range of speeds v to $v + dv$, and thus to a definite chunk of phase space. We denote the associated volume of phase space by $d^3v\, d^3r$. Let Ω denote the volume of each phase-space cell. The numerical value of Ω is of no consequence at the moment; it conveniently cancels out of most of the ensuing computations. We can express $dQ(\varepsilon)$ as

$$\begin{aligned} dQ(\varepsilon) &\equiv \text{number of phase cells in range } \varepsilon \text{ to } \varepsilon + d\varepsilon \\ &= \frac{\text{corresponding volume of phase space}}{\text{volume of one cell}} \\ &= \frac{d^3v\, d^3r}{\Omega} \end{aligned} \tag{15.41}$$

(15.39) becomes

$$Z(\beta) = \frac{1}{\Omega}\int_r d^3r \int_v \exp\left(\frac{-\beta m v^2}{2}\right) d^3v \tag{15.42}$$

The integrand is independent of position so that the integration over position space results in a factor V, the container volume, leaving

$$Z(\beta) = \frac{V}{\Omega}\int_v \exp\left(\frac{-\beta m v^2}{2}\right) d^3v \tag{15.43}$$

The spherical symmetry of the integrand allows us to integrate over spherical shells. Taking $d^3v = 4\pi v^2\, dv$ results in

$$4\pi \int_{v=0}^{\infty} \exp\left(\frac{-\beta m v^2}{2}\right) v^2\, dv = \left(\frac{2\pi}{m\beta}\right)^{3/2} \tag{15.44}$$

The partition function is then

$$Z(\beta) = \frac{V}{\Omega}\left(\frac{2\pi}{m\beta}\right)^{3/2} \equiv KV\beta^{-3/2} \tag{15.45}$$

Although the unit cell volume Ω appears in $Z(\beta)$ it drops out of the calculations which follow. The internal energy per atom now follows at once from (15.45)

$$\frac{U}{N} = -\frac{\partial}{\partial\beta}\left(-\frac{3}{2}\ln\beta - \ln V - \ln K\right) = \frac{3}{2}\frac{1}{\beta}$$

Thus as advertised previously

$$\frac{U}{N} = \frac{3}{2}\frac{1}{\beta} \tag{15.37}$$

However, the kinetic theory model of the ideal gas led to the result $U/N = (3/2)kT$, from which it follows that

$$\beta = \frac{1}{kT} \tag{15.27}$$

This establishes that $\beta = 1/kT$ for an ideal gas. However, $\beta = 1/kT$ in all cases where the Boltzmann distribution prevails. A proof of this statement is presented in Appendix B. The proof proceeds by showing that β has the same value for two systems in mutual thermodynamic equilibrium, regardless of their physical nature. Allowing one of the systems to be an ideal gas, together with the thermodynamic requirement that both systems have the same temperature establishes that $\beta = 1/kT$ for all systems.

....

15.8.2 A system with f degrees of freedom obeys the equipartition of energy theorem. Using the fact that

$$\frac{U}{N} = -\frac{\partial}{\partial\beta}\ln Z(\beta)$$

show that the partition function must be proportional to $T^{f/2}$.

....

We note in passing that, when energy is treated as a continuous variable, the discrete version of the Boltzmann distribution

$$N_i = \frac{NQ_i \exp(-\beta\varepsilon_i)}{Z(\beta)} \tag{15.33'}$$

is replaced by

$$dN(\varepsilon) = \frac{N\,dQ(\varepsilon)}{Z(\beta)} e^{-\beta\varepsilon} \tag{15.33a}$$

The quantity

$$\frac{dN(\varepsilon)}{dQ(\varepsilon)} = \frac{Ne^{-\beta\varepsilon}}{Z(\beta)} \tag{15.33b}$$

is the number of atoms per cell for phase-space cells corresponding to energies in the range ε to $\varepsilon + d\varepsilon$.

....

15.8.3 The expression for the partition function of an ideal monatomic gas given by (15.39) may also be written as

$$Z(\beta) = \int_{\varepsilon=0}^{\infty} e^{-\beta\varepsilon}\rho(\varepsilon)\,d\varepsilon$$

where $\rho(\varepsilon)$ gives the density of cells along the energy axis and is generally referred to as the density of states. The density of states is defined by the statement that $\rho(\varepsilon)\,d\varepsilon$ equals the number of phase-space cells with energies in the range ε to $\varepsilon + d\varepsilon$. Start from (15.43), use the relations $\varepsilon = \frac{1}{2}mv^2$ and $d^3v = 4\pi v^2\,dv$, and show that

$$Z(\beta) = \frac{4\pi V}{\Omega}\sqrt{\frac{2}{m^3}}\int_0^{\infty} e^{-\beta\varepsilon}\sqrt{\varepsilon}\,d\varepsilon$$

that is, that

$$\rho(\varepsilon) = \frac{2\pi V}{\Omega}\left(\frac{2}{m}\right)^{3/2}\sqrt{\varepsilon}$$

Evaluate the integral and verify that (15.45) again results. [*Hint:* It is necessary to form the differential of the relation $\varepsilon = \frac{1}{2}mv^2$ in order to relate $d\varepsilon$ and dv.]

....

Example 15.2 Maxwellian Velocity Distribution Function

Chapter 12 promised a derivation of the Maxwellian velocity distribution law. With the results of Section 15.8 fresh at hand we can quickly fulfill that promise. From (15.33a) we have

$$dN(\varepsilon) = \frac{Ne^{-\beta\varepsilon}}{Z(\beta)} dQ(\varepsilon) \tag{15.33a}$$

as the number of atoms with energies in the range ε to $\varepsilon + d\varepsilon$. Next we consider a group of

$$dQ = \frac{V d^3v}{\Omega}$$

cells, each of whose energies differs only microscopically from $\varepsilon = \frac{1}{2}mv^2$.† Inserting the expression just obtained for $Z(\beta)$ gives (note that Ω cancels out)

$$dN = N\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(\frac{-\beta mv^2}{2}\right) d^3v \tag{15.46}$$

The distribution function $f(v)$ is defined as the density of particles in velocity space

$$dN = f(v)\, d^3v \tag{12.3'}$$

Comparing (15.46) and (12.3') shows that

$$f(v) = N\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(\frac{-mv^2}{2kT}\right) \tag{12.5}$$

the promised Maxwellian distribution. ◇

15.9 THE STATISTICAL INTERPRETATION OF ENTROPY

Consider two systems, A and B, initially isolated and each separately in thermodynamic equilibrium at the same temperature T. We may suppose that the two systems are initially separated by a rigid adiabatic wall. Collectively A and B form a "super system" AB which is isolated and in thermal equilibrium.

† We have taken the coordinate space–volume element equal to the container volume ($d^3r = V$). Thus dQ refers to the total number of cells corresponding to velocities in d^3v.

Let $W(\mathrm{A})$ and $W(\mathrm{B})$ denote the thermodynamic probabilities for A and B, with $W(\mathrm{A})$ the number of macrostates which correspond to the thermodynamic state of A; $W(\mathrm{B})$ the number of macrostates which correspond to the thermodynamic state of B. With A and B isolated, the thermodynamic probability for the super system $W(\mathrm{AB})$ is simply the product of $W(\mathrm{A})$ and $W(\mathrm{B})$

$$W(\mathrm{AB}) = W(\mathrm{A}) \cdot W(\mathrm{B}) \tag{15.47}$$

This relation follows because A and B are isolated. They can independently exist in $W(\mathrm{A})$ and $W(\mathrm{B})$ macrostates so that $W(\mathrm{A}) \cdot W(\mathrm{B})$ macrostates are possible for the combined system.

Suppose that we now strip away the adiabatic coating from the wall separating A and B. No changes occur in the thermodynamic states of A or B because they have the same temperature. The state of the super system is likewise unchanged. It follows that the thermodynamic probability of the super system is still given by (15.47). The act of establishing thermal contact does not alter the number of macrostates open to A and B if they already have equal temperatures. Thus (15.47) holds whenever A and B are in mutual thermal equilibrium.

Next, we note that the extensive properties of the super system, its mass, internal energy, entropy, and so on, are additive. Thus

$$M(\mathrm{AB}) = M(\mathrm{A}) + M(\mathrm{B}) \tag{15.48}$$

$$U(\mathrm{AB}) = U(\mathrm{A}) + U(\mathrm{B}) \tag{15.49}$$

$$S(\mathrm{AB}) = S(\mathrm{A}) + S(\mathrm{B}) \tag{15.50}$$

where M, U, and S denote mass, internal energy, and entropy, respectively. It is (15.50) which concerns us here. The entropy of a thermodynamic state is somehow related to the thermodynamic probability for that state. The functional form of the relation between entropy and thermodynamic probability must be such that both (15.47) and (15.50) are satisfied. It is readily verified that taking

$$S = C \ln W, \qquad C = \text{constant} \tag{15.51}$$

satisfies the prescribed conditions. Thus if we apply (15.51) to the super system and use (15.47), we find

$$S(\mathrm{AB}) = C \ln W(\mathrm{AB}) = C \ln W(\mathrm{A}) + C \ln W(\mathrm{B})$$

However, $C \ln W(\mathrm{A}) = S(\mathrm{A})$, and $C \ln W(\mathrm{B}) = S(\mathrm{B})$; it follows that

$$S(\mathrm{AB}) = S(\mathrm{A}) + S(\mathrm{B}) \tag{15.50}$$

Thus a logarithmic relation leads to the required additive (extensive) property for entropy. The constant C is the same for all systems and must have the

dimensions of entropy but is otherwise arbitrary at this point. We may fix its value by evaluating $\ln W$ for a specific system whose entropy is already known. The particular system which springs to mind is the ideal monatomic gas (I.G.) whose entropy is (for one mole)

$$S_{\text{I.G.}} = R(\ln V + \tfrac{3}{2}\ln T + \text{constant}) \tag{6.16'}$$

We must now evaluate $\ln W$, where

$$W = N! \prod_{i=1}^{M} \frac{Q_i^{N_i}}{N_i!} \tag{15.19'}$$

and where the most probable population of cells is given by the Boltzmann distribution

$$\frac{N_i}{Q_i} = \frac{N \exp(-\beta\varepsilon_i)}{Z(\beta)} \tag{15.33'}$$

Recall first that the logarithm of a product equals the sum of the logarithms of the factors. Thus

$$\ln W = \ln N! + \sum_{i=1}^{M} (N_i \ln Q_i - \ln N_i!)$$

Next we use the super Stirling approximation for $\ln N!$ and for each of the $\ln N_i!$. This gives

$$\ln N! = N \ln N - N$$

and

$$-\sum_{i=1}^{M} \ln N_i! = -\sum_{i=1}^{M} (N_i \ln N_i - N_i)$$

Taking account of the fact that $N = \sum_{i=1}^{M} N_i$, and that

$$\sum_{i=1}^{M} (N_i \ln Q_i - N_i \ln N_i) = -\sum_{i=1}^{M} N_i \ln \frac{N_i}{Q_i}$$

leaves us with

$$\ln W = N \ln N - \sum_{i=1}^{M} N_i \ln \frac{N_i}{Q_i}$$

For the most probable distribution (15.33)

$$\ln \frac{N_i}{Q_i} = \ln N - \ln Z(\beta) - \beta\varepsilon_i$$

Once again taking account of the fact that $N = \sum_{i=1}^{M} N_i$ shows that the $N \ln N$ term cancels, leaving

$$\ln W = N \ln Z(\beta) + \sum_{i=1}^{M} N_i \varepsilon_i \tag{15.52}$$

We recognize $\sum_{i=1}^{M} N_i \varepsilon_i$ as the internal energy U. Thus

$$\ln W = N \ln Z(\beta) + \beta U \tag{15.53}$$

It should be noted that this result for $\ln W$ is completely† general and is not restricted to the ideal gas. To evaluate $\ln W$ for the ideal gas we use (15.37) to find

$$\beta U = \tfrac{3}{2} N \qquad \text{(ideal gas)}$$

while (15.45) gives

$$N \ln Z(\beta) = N(\ln V + \tfrac{3}{2} \ln T + \text{constant})$$

Inserting these results into (15.53) and taking $N = N_0$ (since we deal with 1 mole of gas) gives

$$(C \ln W)_{\text{I.G.}} = CN_0(\ln V + \tfrac{3}{2} \ln T + \text{constant}) \tag{15.51a}$$

Comparing this with (6.16′) shows that $CN_0 = R$ or

$$C = \frac{R}{N_0} = k = 1.38 \simeq 10^{-16} \quad \frac{\text{erg}}{\text{K}} \tag{15.54}$$

Thus $C = k$, the Boltzmann constant, whereupon (15.51) becomes

$$S = k \ln W \tag{15.55}$$

This quantitative relation between entropy and thermodynamic probability unites two previous pronouncements concerning equilibrium, one thermodynamic the other statistical. To wit,

The state of maximum entropy corresponds to the thermodynamic equilibrium state;

and

The most probable macrostate corresponds to the thermodynamic equilibrium state.

The relation $S = k \ln W$ establishes the connection between entropy, a macroscopic quantity, and thermodynamic probability, a microscopic concept. In Chapter 6 we observed that an entropy increase was the result of processes which disordered or randomized a system. Our observations there helped to

† Provided we restrict ourselves to systems which are adequately described by classical physics.

establish a valuable but purely qualitative feeling for entropy. The key shortcoming was the lack of any quantitative definition of disorder. The relation $S = k \ln W$ provides the quantitative link between the concepts of entropy and disorder. Disorder is measured by $\ln W$. At least it is plausible and consistent to interpret $\ln W$ as a direct measure of disorder. For example, a highly-ordered distribution of atoms can be formed in relatively few ways, and therefore has a low thermodynamic probability, whereas the most probable distribution can be formed in the largest number of ways. The most probable distribution gives the greatest randomization, or mixing, consistent with the constraints.

....

15.9.1 Show that

$$S(xy) = S(x) + S(y) \tag{15.50'}$$

has the solution

$$S(x) = C \ln x$$

by forming the partial derivatives of $S(xy)$ with respect to x and y. Proceed as follows: Observe that, with $z \equiv xy$, the chain rule leads to

$$\frac{\partial S(z)}{\partial x} = y \frac{\partial S(z)}{\partial z} \quad \text{and} \quad \frac{\partial S(z)}{\partial y} = x \frac{\partial S(z)}{\partial z}$$

so that

$$\frac{1}{y} \frac{\partial S(xy)}{\partial x} = \frac{1}{x} \frac{\partial S(xy)}{\partial y}$$

Having established this result, use (15.50) to show

$$h(x) \equiv x \frac{\partial S(x)}{\partial x} = y \frac{\partial S(y)}{\partial y} = h(y)$$

and remember that if $h(x) = h(y)$ for all values of two independent variables x and y, $h(x) = h(y) =$ constant.

....

15.10 THERMODYNAMIC PROPERTIES: RELATIONSHIP TO THE PARTITION FUNCTION

In addition to supplying a conceptual link, (15.55) provides the basis for relating the partition function to the thermodynamic properties of a system. Using (15.53) in (15.55) gives a general relation between S, U, and $Z(\beta)$

$$S = k\beta U + kN \ln Z(\beta) \tag{15.56}$$

Multiplying through by T and noting that $kT\beta = 1$ results in

$$U - TS = -NkT \ln Z(\beta) \tag{15.57}$$

However, $U - TS$ is the Helmholtz free energy F, whence,

$$F = -NkT \ln Z(\beta) \tag{15.58}$$

We have previously established the relationship between $Z(\beta)$ and the internal energy U [compare (15.36)]. The Gibbs free energy G and the enthalpy H also may be expressed in terms of $Z(\beta)$.

....

15.10.1 Verify that the Gibbs free energy G and the enthalpy H are related to the partition function by

$$\frac{G}{N} = kT\left[V\frac{\partial}{\partial V}\ln Z - \ln Z\right]$$

$$\frac{H}{N} = \frac{V}{\beta}\frac{\partial}{\partial V}\ln Z - \frac{\partial}{\partial \beta}\ln Z$$

....

From thermodynamics it will be recalled that the equation of state emerges as one of the Maxwell relations

$$P = -\left(\frac{\partial F}{\partial V}\right)_T$$

Using (15.58) it follows that

$$P = \frac{NkT}{Z(\beta)}\left(\frac{\partial Z(\beta)}{\partial V}\right)_\beta \tag{15.59}$$

....

15.10.2 Use (15.59) and the ideal gas partition function to obtain the equation of state

$$P = \frac{NkT}{V}$$

....

The thermodynamic properties of a system, its internal energy, entropy, and so on, and its equation of state all follow from the partition function. To know $Z(\beta)$ is to know all at the thermodynamic level. There are enormous mathematical difficulties in evaluating $Z(\beta)$ for realistic models of specific physical

systems. Failure to obtain agreement with experiment may be traced to some deficiency of the model because, as we have noted previously, our faith in the principles of statistical mechanics rests primarily on one supreme fact; the laws of thermodynamics may be derived from statistical mechanics. Of course the fun is in playing the game, that is, in applying statistical mechanics to specific physical systems. Proving that thermodynamics follows from statistical mechanics involves too much mathematical juggling to be classified as fun. The student interested in juggling should consult the references. The next chapter is devoted to the fun aspects of statistical mechanics.

15.11 SUMMARY

Several key concepts have been developed in this chapter. Among these are the notions of a microstate and a macrostate. A *microstate* specifies which atoms occupy each phase-space cell. A *macrostate* tells only how many atoms occupy each group of cells.

Our disbelief in Maxwell demons, that is, our refusal to accept the possibility that macroscopic constraints can control atomic motions, led to the postulate of equal *a priori* probability for all accessible microstates.

The *thermodynamic probability* was introduced, being defined as the number of microstates corresponding to the same macrostate. The most probable macrostate is the one having the largest thermodynamic probability.

The fact that thermodynamic equilibrium persists while a system moves through a succession of different macrostates was noted. This suggests that many different macrostates are equally valid representatives of the gross (macroscopic) behavior of a system. It proves mathematically feasible to determine the nature of the most probable macrostate. We therefore settle on it as *the* representative macrostate. That is, we assume that the thermodynamic equilibrium state has the same macroscopic properties as the most probable macrostate.

The most probable macrostate results when the members of the system have a Boltzmann distribution with respect to energy

$$N_i = \frac{NQ_i \exp\left(\frac{-\varepsilon_i}{kT}\right)}{Z(\beta)}$$

where $Z(\beta)$ is the partition function

$$Z(\beta) = \sum_{i=1}^{M} Q_i \exp(-\beta\varepsilon_i), \qquad \beta \equiv \frac{1}{kT}$$

The summation runs over groups of cells with different energies, Q_i being the number of cells with energy ε_i, and M denotes the total number of groups of cells.

The Boltzmann distribution reveals the "competition" for energy. The energy characteristic of the thermal motion of a particle is kT. Phase-space cells corresponding to energies much greater than kT are sparsely populated.

The entropy of a system is related to the thermodynamic probability by (15.55)

$$S = k \ln W$$

This relation reinforces our earlier observations which suggested that entropy measures the randomization or disorder of a system.

Finally, we have seen that the partition function $Z(\beta)$ is a key quantity, leading directly to the four thermodynamic potentials U, F, G, and H.

REFERENCES

Recommended for its clarity and wit is:

K. K. Darrow, Memorial to the Classical Statistics, *Bell Sys. Tech. J.* **22**, 108 (1943).

Most texts derive the Boltzmann distribution by using the method of Lagrangian multipliers. Methods intermediate between the one employed in Section 15.7 and the full Lagrangian treatment are given in the following three references:

R. Weinstock, On Introducing the Boltzmann Distribution, *Contemporary Phys.* **7**, 234 (1966).

D. Kleppner, Avoiding Lagrange Multipliers in Introductory Statistical Mechanics, *Am. J. Phys.* **36**, 843 (1968).

J. W. Lorimer, Elementary Statistical Mechanics without Lagrange Multipliers, *J. Chem. Educ.* **43**, 39 (1966).

A concise and coherent discussion of microstates, macrostates, and thermodynamic probability is given in:

F. W. Sears, "An Introduction to Thermodynamics, The Kinetic Theory of Gases, and Statistical Mechanics," 2nd ed., Chapter 14. Addison-Wesley, Reading, Massachusetts (1953).

Chapter 16

Applications of Statistical Mechanics

In this chapter the methods of statistical mechanics are applied to three rather diverse topics. At places the discussion is just a notch above the qualitative level. The aim of the chapter is threefold; to impress the student with the scope and power of statistical mechanics, to reinforce the ideas and formalism presented in the preceding chapter, and to pave the way for the presentation of quantum statistical mechanics in the next chapter.

In Section 16.1 the microscopic basis for the Clausius–Clapeyron equation is revealed. In Section 16.2 several aspects of chemical reactions are studied from the viewpoint of statistical mechanics. The chemist's "ten-degree rule"—which states that the rate of a reaction is doubled by a ten-degree increase in temperature—is shown to have a statistical mechanical basis. Finally, in Section 16.3 the statistical character of paramagnetism is studied and the origin of Curie's law uncovered.

16.1 MICROSCOPIC BASIS OF THE CLAUSIUS–CLAPEYRON EQUATION

The Clausius–Clapeyron equation was derived in an earlier chapter by considering a special sort of Carnot cycle. For a liquid–vapor transition, the slope of the vapor pressure versus temperature curve is given by the

Clausius–Clapeyron equation in the form

$$\frac{dP_v}{dT} = \frac{L_v}{T(V_v - V_l)} \simeq \frac{L_v}{TV_v} \tag{16.1}$$

where P_v is the vapor pressure, T is the absolute temperature, L_v is the molar latent heat of vaporization, and V_v and V_l are the molar volumes of vapor and liquid. Because of the large expansion in the liquid-to-vapor transition $V_v \gg V_l$, and the approximation in (16.1) is generally quite satisfactory. We now want to see inside the Clausius–Clapeyron equation—to understand it at the molecular level. Statistical mechanics can furnish such insight.

Consider a container, partially filled with liquid, and let N_v and N_l denote the numbers of molecules in equal volumes of vapor and liquid (see Figure 16.1). If the vapor is not too dense, the molecules in it will behave like those of

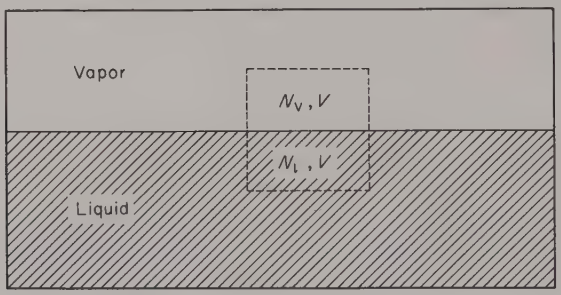


Figure 16.1. The liquid and vapor are in equilibrium at temperature T.

an ideal gas. The characteristic energy of vapor molecules is then $\frac{1}{2}kT$ per degree of freedom. The molecules in the liquid fall much lower on the energy scale. Strong intermolecular forces cause them to cling together in a configuration whose average energy per molecule is largely in the form of potential energy. This potential energy is negative and its magnitude is much greater than the mean kinetic energy of molecular motion. For these reasons we can write for the energies characterizing the vapor and liquid states as

$$\varepsilon_v \simeq kT$$

$$\varepsilon_l \simeq -E, \qquad E \gg kT$$

If we now regard the vapor and liquid as two different equilibrium states of a system characterized by energies ε_v and ε_l, the Boltzmann distribution law sets the ratio of populations at

$$\frac{N_v}{N_l} = \frac{\exp(-\varepsilon_v/kT)}{\exp(-\varepsilon_l/kT)} \tag{16.2}$$

Inserting the energies representative of liquid and vapor and noting that $E \gg kT$ produces

$$\frac{N_v}{N_l} \simeq e^{-E/kT} \tag{16.3}$$

The competition for thermal energy, as described by the Boltzmann distribution, is the origin of the exponential temperature dependence of the vapor pressure which we found in Chapter 5. Equation (16.3) leads directly to the Clausius–Clapeyron equation and in doing so adds support to the plausible notion that the latent heat of vaporization is the energy necessary to break the intermolecular bonds and thereby convert one mole of liquid into one mole of vapor. It must be confessed that (16.3) is not a rigorous application of the Boltzmann formula because we have used the average energies of two groups of particles whose individual energies are spread over fairly wide ranges. However, because the average energies differ so markedly, no essential error results if we consider the vapor and liquid as two different energy states of a single system. Multiplying both sides of (16.3) by $N_l kT/V$ gives

$$\frac{N_v kT}{V} = \frac{N_l kT}{V} e^{-E/kT} \tag{16.4}$$

If we recall the ideal gas equation of state in the form $P = NkT/V$, we recognize the left side as the vapor pressure. Thus

$$P_v = \frac{N_l kT}{V} e^{-E/kT} \tag{16.5}$$

The factor N_l/V is the number density for the liquid, a quantity which is virtually independent of temperature. To convert this into the Clausius–Clapeyron form we differentiate P_v with respect to T, ignoring any slight temperature dependence of the liquid density. The result is

$$\frac{dP_v}{dT} = \frac{P_v}{T}\left(1 + \frac{E}{kT}\right)$$

In view of the fact that $E \gg kT$ we can take this to be

$$\frac{dP_v}{dT} = \frac{P_v}{T}\frac{E}{kT}$$

Finally, if we note again that $P_v/kT = N_v/V$, we have

$$\frac{dP_v}{dT} = \frac{N_v E}{TV} \tag{16.6}$$

To compare this with the Clausius–Clapeyron equation, let N_v and V refer to 1 mole of vapor

$$N_v = N_0, \quad \text{Avogadro's number}$$

$$V = V_v, \quad \text{molar volume of vapor}$$

Then

$$\frac{dP_v}{dT} = \frac{N_0 E}{TV_v} \tag{16.7}$$

which concurs with the Clausius–Clapeyron equation (16.1) provided we identify $N_0 E$ with L_v, the latent heat. This is precisely how we should interpret $N_0 E$, for we originally argued that E was the difference between the average energies of a molecule in the vapor and liquid phase. Thus $N_0 E$ is the difference in energy per mole—the energy we must add to 1 mole of liquid to convert it into vapor at the same temperature. With

$$N_0 E = L_v$$

(16.7) assumes the form of the Clausius–Clapeyron equation.

....

16.1.1 The relation $N_0 E = L_v$ enables us to check the validity of the assumption $E \gg kT$.

(a) Determine the ratio E/kT_B for the liquids listed in the accompanying tabulation. (T_B is the normal boiling point.) Note that L_v is expressed in calories per mole.

(b) Evaluate E and kT_B for water and hydrogen. Express the energies in electron volts.

Substance	L_v (cal/mole)	T_B (K)
Water	9717	373
Ethyl ether	6870	308
Ammonia	5560	240
Hydrogen sulfide	4490	213
Oxygen	1610	90
Hydrogen	216	20

....

From (16.5) it is evident that the vapor pressure falls quite rapidly with temperature. From our inside vantage point we can interpret this as a direct consequence of the exponential Boltzmann factor, which favors states of low energy. In this instance it is the liquid state with its negative potential energy

configurations which is favored over the vapor state in which the molecules have modest, but positive energies.

The Boltzmann factor is in evidence whenever there is competition for energy at the atomic-molecular level—it is the recurring theme of statistical physics.

16.2 CHEMICAL REACTIONS

In this section three aspects of chemical reactions are discussed from the statistical mechanical viewpoint. First the temperature dependence of chemical equilibrium constants is investigated. This is followed by a discussion of activation energies. Finally, an empirical rule of thumb—the ten-degree rule—is shown to have a statistical mechanical basis.

A Equilibrium Constants

We select as a model reaction the formation and dissociation of a diatomic molecule AB from atoms A and B

$$A + B \rightleftharpoons AB \tag{16.8}$$

The formation of an AB molecule occurs whenever A and B atoms collide under favorable conditions. Dissociation results when AB molecules acquire energy enough to tear themselves apart. Figure 16.2 suggests the energy level scheme for the model reaction. The energy E is the dissociation energy of the

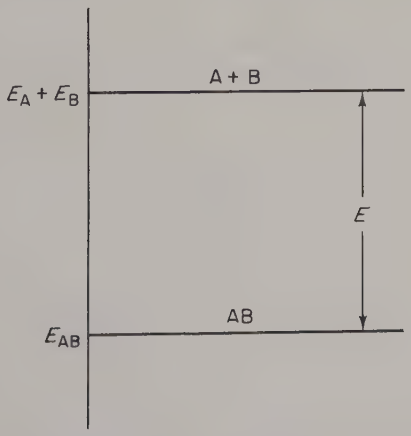


Figure 16.2. An energy level scheme for the reaction $A + B \rightleftharpoons AB$. Adding energy E to the molecule AB causes its dissociation into the atoms A and B.

molecule, the minimum energy which must be acquired by the molecule in order for it to break up into its component atoms A and B. The value of E is a characteristic of the particular reaction under study, not some universal constant. However, many such energies fall in the electron volt (eV) range

$$1 \text{ eV} = 1.6 \times 10^{-12} \text{ erg} = 3.8 \times 10^{-20} \text{ cal}$$

....

16.2.1 The Coulomb potential energy of two charges, q_1 and q_2, separated by a distance r is $q_1 q_2/r$. The characteristic atomic charge is the electronic charge, $e = 4.8 \times 10^{-10}$ esu, and the characteristic atomic dimension is 1 Å = 10^{-8} cm.

(a) Evaluate the corresponding characteristic Coulomb energy in electron volts and calories (1 erg = 1 esu^2/cm).

(b) Can you suggest why molecular dissociation energies are smaller than the corresponding atomic ionization energies? (To cite a typical example, it requires about 4.7 eV to dissociate the hydrogen molecule, but nearly 13.6 eV to ionize the hydrogen atom.)

....

Let n_A, n_B, n_{AB} denote the concentrations (number densities) of the reaction participants. If n_A were doubled, the probability of "A meets B" would be doubled and hence the rate of molecule-forming collisions would also be doubled. Likewise, doubling n_B doubles the chance for an encounter between A and B atoms thereby doubling the rate of AB formation. The rate of AB production is therefore proportional to the product of the concentrations $n_A n_B$. The rate of formation is also proportional to the mean relative speed at which A and B atoms encounter each other, which in turn, is proportional to the square root of the absolute temperature. However, there are other temperature-dependent factors affecting the rate of formation. These additional factors depend strongly on the nature of A and B as well as the temperature; they are not universal factors. If we lump together all of the temperature-dependent factors into a reaction constant $K_f(T)$ we can write the rate of formation as

$$R_f = K_f(T) n_A n_B \tag{16.9}$$

The rate of dissociation depends on three quantities: (a) the number of AB molecules present, (b) the probability that a molecule has acquired enough energy to dissociate, and (c) the probability that dissociation occurs, given that the needed energy is available. Only if an AB molecule has sufficient energy—at least E—can dissociation proceed. The competition for energy does not favor dissociation. The probability that a molecule has acquired the dissociation energy E is proportional to the Boltzmann factor $e^{-E/kT}$. The total number

of molecules with energy enough to dissociate is therefore proportional to $n_{AB}e^{-E/kT}$, and the rate of dissociation can be written as

$$R_d = K_d(T)n_{AB}e^{-E/kT} \tag{16.10}$$

Our ignorance of the third factor cited above is buried in the dissociation constant $K_d(T)$, which, like $K_f(T)$, depends on the specific nature of the reaction.

Chemical equilibrium prevails when the rates of formation and dissociation are equal, $R_f = R_d$. From (16.8) and (16.10) it follows that, at equilibrium,

$$\frac{n_A n_B}{n_{AB}} = \frac{K_d(T)}{K_f(T)}e^{-E/kT} \equiv K(T) \tag{16.11}$$

The form of the ratio $K_d(T)/K_f(T)$ depends strongly on the details of the chemical reaction. However, as we shall see, the crucial temperature dependence of the equilibrium constant $K(T)$ resides in the Boltzmann factor, which is universally present,

$$K(T) = \frac{K_d(T)}{K_f(T)}e^{-E/kT} \tag{16.12}$$

Let us summarize the development to this point; the equilibrium constant $K(T)$ relates the equilibrium concentrations n_A, n_B, and n_{AB} of the principals in the reaction

$$A + B \rightleftharpoons AB$$

By equating the rates of molecular formation and dissociation at equilibrium we arrived at (16.12), showing that the equilibrium constant $K(T)$ has the form of a product. One factor, the ratio $K_d(T)/K_f(T)$, has a gradual or weak temperature dependence. The decisive importance of the universal Boltzmann factor $e^{-E/kT}$ remains to be demonstrated (compare, however, Problem 16.2.3).

....

16.2.2 In the molecular reaction

$$AB + CD \rightleftharpoons AC + BD$$

an amount of energy E is released (per reaction) when AB and CD combine to give AC and BD.

(a) Show that the equilibrium concentrations of the four constituents are related by an equation of the form

$$K(T) = \frac{n_{AB}n_{CD}}{n_{AC}n_{BD}}$$

(b) $K(T)$ contains an exponential Boltzmann factor. Explain its origin.

16.2.3 Consider the function

$$K(x) = xe^{-30/x}$$

over the interval from $x = 3$ to $x = 7$. Tabulate or plot $K(x)$ and the approximations

$$K_1(x) = \langle x \rangle e^{-30/x}$$

$$K_2(x) = xe^{-30/\langle x \rangle}$$

where $\langle x \rangle$ is the arithmetic average of x over the interval. Which approximation is the more faithful?

....

B Activation Energies

With reference to our model reaction

$$A + B \rightleftharpoons AB$$

the formation of a stable molecule is said to be energetically possible, or favored. By forming a molecule, the atoms decrease their mutual potential energy. The energy released shows up largely as the heat of reaction. Frequently situations arise wherein an energetically possible reaction does not go or else proceeds at an extremely slow pace. The formation of water vapor according to the scheme

$$2H_2 + O_2 \rightleftharpoons 2H_2O$$

is a well-known example. If hydrogen and oxygen are mixed at ordinary temperatures, the above reaction proceeds so slowly as to be imperceptible. However, if an electric spark or a suitable catalyst is introduced, the reaction rate can jump to explosive proportions. Similar situations occur frequently in organic reactions. Often it is necessary for the reactants to form an intricate geometrical configuration—an *activated complex* as it is called—before they can reach a final product form. A catalyst speeds up the reaction by aiding the formation of activated complexes. For example, enzymes apparently serve only as templates in biological reactions. They help the large and clumsy biological molecules maneuver into activated complexes.

In order to form an activated complex the reactants must acquire a certain activation energy. The catalyst loans energy to the reactants and is repaid as the complex breaks up to form the reaction products. Mechanically the situation is analogous to the one shown in Figure 16.3a. In order to reach the position of lower potential energy the skier must first increase his potential energy by surmounting the hill. He accomplishes this by doing an amount of work mgH against the force of gravity. Energetically, the activated complex is

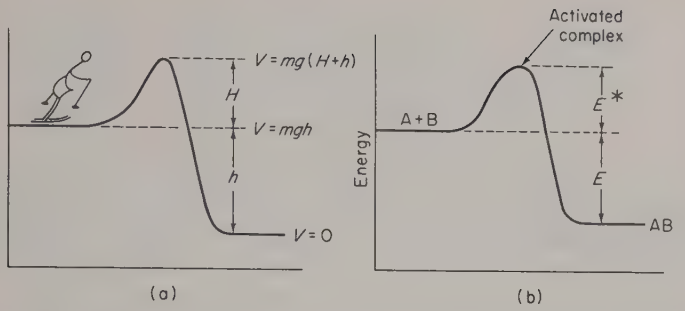


Figure 16.3. (a) To reach the lower level, the skier must first reach the hillcrest; this, in turn, requires that he increase his potential energy. (b) The activation energy E^* is the energy needed by A and B to form an activated complex.

a hilltop, as suggested by Figure 16.3b. The reacting atoms A and B must acquire an activation energy E^* in order to form a complex from which they can proceed to the final product state. In many cases the activation energy is much larger than the characteristic thermal energy kT. Only rarely then will the kinetic energies of the reacting atoms be as large as E^*, and the reaction rate will be small unless a catalyst lends a hand—and some energy.

Our earlier analysis is readily modified to accommodate the effects of an activation energy. The formation rate given by (16.9) must be altered by multiplying the right side by the probability that A and B possess the activation energy E^*. This probability is proportional to the Boltzmann factor $e^{-E^*/kT}$. Thus the rate of formation can be written as

$$R_f = K_f n_A n_B e^{-E^*/kT}$$

The rate of dissociation is given by (16.10), but with $E+E^*$ replacing E. For an AB molecule to break up it must first revert to the activated complex form and this demands an energy $E+E^*$

$$R_d = K_d n_{AB} e^{-(E+E^*)/kT}$$

Equating the rates of formation and dissociation shows that the equilibrium concentrations are still related by (16.11) since the common factor $e^{-E^*/kT}$ cancels. So, the existence of an activation energy has no bearing on the equilibrium concentrations of the constituents. It can dramatically affect the *rate* at which equilibrium is reached. For example, suppose that E^* is 1 eV and that T is 400 K. Then kT is roughly (1/30) eV and

$$e^{-E^*/kT} = e^{-30} \simeq 10^{-13}$$

Under these circumstances the rate of formation is 10^{-13} below what it would be in the absence of any activation energy demands—a dramatic reduction indeed.

.....

16.2.4 Assume a Maxwellian distribution of atoms at temperature T and determine the fraction with kinetic energies in excess of $30kT$. [*Hint:* Refer to Section 12.5 and to Figure 12.6 in particular.]

.....

In the final segment of this section we investigate how the reaction rate changes with temperature; in doing so we uncover the basis for the chemist's ten-degree rule.

C The Ten-Degree Rule

Chemists and biologists have long used a rule of thumb which states that the rate of a chemical reaction doubles with a 10 °C increase in temperature. The figure of 10 °C is a rough approximate value; the precise change in temperature required to double the rate of a reaction depends on the nature of the compounds involved and on the temperature itself. The basis for the rule stems from the fact that the characteristic activation energies associated with a great many chemical reactions are of roughly the same magnitude—namely, about one electron volt (1 eV) on a per-molecule basis.

In the chemical reaction

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$$

the rate at which A and B atoms unite to form AB molecules can be expressed as

$$R = K_{\mathrm{f}} n_{\mathrm{A}} n_{\mathrm{B}} e^{-E^*/kT}$$

wherein the decisive temperature dependence of R is contained in the exponential $e^{-E^*/kT}$. At a typical cooking temperature of 400 K, the characteristic activation energy is roughly 1 eV, and E^*/kT is in the neighborhood of 30. As we now demonstrate, a small change in temperature can change $e^{-E^*/kT}$ by a sizeable factor. So—we ignore the generally weak temperature dependence of K_{f} and focus on the exponential. If the temperature changes from T to $T+\delta T$, the exponent changes to

$$\frac{E^*}{kT\left(1+\dfrac{\delta T}{T}\right)} \simeq \frac{E^*}{kT}\left(1-\frac{\delta T}{T}\right)$$

The change in the exponent is seen to be

$$\frac{E^*}{kT}\frac{\delta T}{T} \simeq 30\frac{\delta T}{T}$$

Observe that a 1% change in T gives a 30% change in the exponent. The reaction rate at the temperature $T+\delta T$ may be expressed as

$$R(T+\delta T) = R(T)\exp\left(+\frac{E^*}{kT}\frac{\delta T}{T}\right) \tag{16.13}$$

It is evident how a small fractional change in the temperature can produce a sizable change in the exponent—and thence the reaction rate. The fractional change $\delta T/T$ is multiplied by a whopping $E^*/kT \simeq 30$. To make $R(T+\delta T)$ twice $R(T)$ the exponent in (16.13) must equal $\ln 2 = 0.69$. With $E^*/kT = 30$, and $T = 400$ K, the reaction rate doubles when $\delta T = 9\,°\text{C}$. Here then is the origin of the ten-degree rule. The exponential factor which governs the availability of activation energy is extremely sensitive to slight changes in temperature.

The ten-degree rule has some bearing on the existence of life. A drop in temperature reduces the rates of chemical reactions. Each reaction in the complicated schemes of biological oxidation and photosynthesis is similarly affected. As the temperature falls, the reaction rates diminish until life becomes dormant or ceases.

A healthy warm-blooded animal maintains a constant body temperature. A cold-blooded creature assumes the temperature of its surroundings. The built-in thermostats which human beings and other warm-blooded animals enjoy is of the greatest importance. Imagine what life would be like if man were cold-blooded. Going from a heated room at 72 °F to a frigid 0 °F surroundings would result in a 40 °C change of body temperature. The ten-degree rule advises that a 40 ° temperature drop depresses the biological reaction rates by a factor of 16. Something between death and monumental lethargy would surely follow.

A temperature rise increases the activation energy reservoir; a larger fraction of the molecules acquire the needed activation energy. However, two other effects of high temperature prove disastrous for life. As the temperature rises there is an increase in the kinetic energies of the long chain molecules which typify living systems (for example, the DNA molecule). The entire fabric of the living system jiggles and squirms more violently. Some of the molecules literally shake themselves to pieces. More important is the fact that the enzymes, the biological catalysts, gradually lose their shapes. Unable to provide a faithful template for the formation of activated complexes, their biological function ends, and life ceases.

....

16.2.5 The ten-degree rule is only an approximate rule of thumb. The value of the temperature change needed to double the reaction rate depends on the values of E^* and T. Determine the approximate value of δT required to double

the reaction rate when

(a) $E^* = 1$ eV, $T = 300$ K,
(b) $E^* = 2$ eV, $T = 400$ K.

16.2.6 How would measurements of the reaction rate and the temperature permit the activation energy to be determined? [*Hint:* Consider the shape of a $\ln R$ versus $1/kT$ plot of the data.]

....

16.3 THE LANGEVIN THEORY OF PARAMAGNETISM

The response of a magnetic material to an applied magnetic field is measured by the magnetization $\mathcal{M}$, the magnetic dipole moment per unit volume. The total magnetic moment of a substance $\mathbf{M}$ is the vector sum of the moments of individual atoms or ions. An applied magnetic field tends to align individual moments while thermal agitation promotes random orientations. We anticipate that this competition will show itself in the form of an appropriate Boltzmann factor.

Let $\mathbf{H}$ denote the external magnetic field applied, and let $\mathcal{M} = \mathbf{M}/V$ denote the resulting magnetization. The magnetic properties are conveniently classified in terms of the susceptibility χ

$$\mathcal{M} = \chi \mathbf{H} \tag{16.14}$$

The susceptibility of a *diamagnetic* material is negative and temperature independent. The atoms of a diamagnetic substance do not have permanent magnetic moments. The applied field induces a magnetic moment in accord with Lenz's law. That is, the response of the atomic electrons produces a current whose magnetic field opposes the applied field. Diamagnetic effects are swamped by *paramagnetism* in materials whose atoms or molecules have permanent magnetic dipole moments. The dipoles tend to align in a way which augments the applied field. The susceptibility of a paramagnetic material is positive and temperature-dependent. At room temperature χ is roughly 10^{-2}. The paramagnetic susceptibility varies approximately as $1/T$ at ordinary temperatures. Pierre Curie parlayed this fact into what is now known as Curie's law

$$\chi = \frac{C}{T} \tag{16.15}$$

where C, called the Curie constant, is not a universal constant, but is different for each paramagnetic material. Curie's analysis was strictly empirical and his expression for χ is inadequate at a sufficiently low temperature.

In 1905, P. Langevin advanced a theory of paramagnetism. It is essentially

the Langevin theory which we present here. A limiting form of the Langevin theory leads to Curie's law. Our goal is to understand the microscopic basis of the Langevin theory, to derive Curie's law from it, and to speculate as to why it requires modification at low temperatures.

The particles responsible for the paramagnetic behavior of a material may be atoms, or molecules, or even ions. Each such particle has a permanent magnetic dipole moment $\boldsymbol{\mu}$. As a result, each particle behaves like a tiny bar magnet. Because each particle (atom, molecule, ion) for a given material is identical, the size of the dipole moment μ is the same for all. The atomic structure of the particle determines the size of μ. When an external magnetic field is applied the magnetic dipoles experience a torque

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{H} \tag{16.16}$$

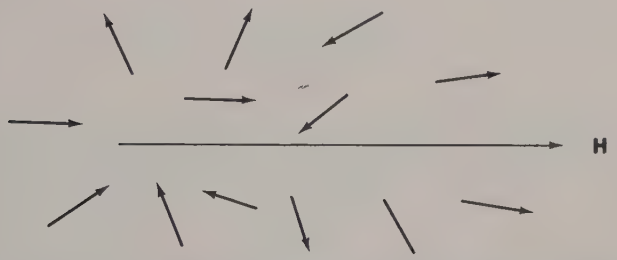


Figure 16.4. The z component of the magnetic dipole moment varies. Orientations with $\boldsymbol{\mu}$ parallel to $\mathbf{H}$ are favored as they correspond to the lowest energy.

This torque tends to rotate the dipole moments into a stable, torque-free position, which is to say, it tends to line up the magnetic moments in the direction of the applied field. Such ordering is only partially successful because of thermal agitation. Thermal motions tend to randomize the orientation of the moments. In the absence of an external field such randomization is complete and the vector sum of the magnetic moments is zero.† So, we have competing effects. Thermal energy tends to randomize; the external field tends to order. Whenever there is a competition for energy the Boltzmann factor appears.

To calculate the magnetization $\mathcal{M}$ and thence the susceptibility χ, we proceed as follows: Take the z axis along the direction of $\mathbf{H}$. The potential energy of a permanent magnetic dipole $\boldsymbol{\mu}$ in the field $\mathbf{H}$ is

$$\varepsilon = -\boldsymbol{\mu} \cdot \mathbf{H} = -\mu_z H \tag{16.17}$$

Although μ is the same for each of the particles the z component of $\boldsymbol{\mu}$ will vary (see Figure 16.4). From (16.17) we see that the favored orientation—the one

† Because individual magnetic moments frequently have their orientation altered it is more precise to say that the vector sum fluctuates about an average value of zero.

of lowest energy—has $\boldsymbol{\mu}$ and $\mathbf{H}$ parallel

$$\varepsilon_{\min} = -\mu H$$

whereas the configuration of maximum energy has $\boldsymbol{\mu}$ antiparallel to $\mathbf{H}$,

$$\varepsilon_{\max} = +\mu H$$

The probability that the z component of the dipole moment lies in the range μ_z to $\mu_z + d\mu_z$ is proportional to $d\mu_z$ and to the Boltzmann factor

$$\exp\left(\frac{-\varepsilon}{kT}\right) = \exp\left(\frac{+\mu_z H}{kT}\right)$$

Thus

$$dP = A\exp(\alpha\mu_z)\,d\mu_z, \qquad \alpha = \frac{H}{kT} \tag{16.18}$$

....

16.3.1 Show that

$$A = \frac{\alpha}{e^{\mu\alpha} - e^{-\mu\alpha}}$$

follows from the requirement that $\int_{-\mu}^{+\mu} dP = 1$. What is the basis for imposing this requirement?

....

The mean value of μ_z is

$$\langle\mu_z\rangle = \int_{-\mu}^{+\mu} \mu_z\, dP$$

The result may be expressed as

$$\langle\mu_z\rangle = \mu\left[\coth(\mu\alpha) - \frac{1}{\mu\alpha}\right] \tag{16.19}$$

where

$$\coth(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}} \tag{16.20}$$

is the hyperbolic cotangent function. The total magnetic moment of a material containing N such dipoles is $N\langle\mu_z\rangle$. The magnetization is just the dipole moment per unit volume $N\langle\mu_z\rangle/V$. Thus

$$\mathcal{M} = \frac{N\mu}{V}\left(\coth(\mu\alpha) - \frac{1}{\mu\alpha}\right) \tag{16.21}$$

The function

$$\mathscr{L}(x) \equiv \coth(x) - \frac{1}{x} \tag{16.22}$$

is termed the *Langevin function*. Under ordinary conditions, the variable

$$x = \mu\alpha = \frac{\mu H}{kT} \tag{16.23}$$

is generally much smaller than unity. For example, the magnetic moment has a magnitude of roughly one Bohr magneton† μ_B

$$\mu_B \equiv \frac{e\hbar}{2mc} = 0.927 \times 10^{-20} \quad \frac{\text{erg}}{\text{G}} \tag{16.24}$$

With $H = 1000$ G and $T = 300$ K,

$$\frac{\mu_B H}{kT} \simeq 2 \times 10^{-4}$$

Note that x is the ratio of two characteristic energies; μH is the energy associated with the magnetic ordering, and kT is the characteristic thermal energy. The fact that $x = \mu H/kT$ is small at ordinary temperatures means that the magnetic ordering is weak or incomplete.

To obtain Curie's law we make use of the fact that x is small and use an approximate form of the Langevin function (see Problem 16.3.2)

$$\mathscr{L}(x) \simeq \frac{x}{3}, \qquad x \ll 1$$

Using this approximation the Langevin function in (16.21) gives

$$\mathscr{M} = \frac{N\mu^2 H}{3VkT} \tag{16.25}$$

Since $\mathscr{M} = \chi H$ defines the susceptibility we have

$$\chi = \frac{N\mu^2}{3VkT} \tag{16.26}$$

which is Curie's law, the Curie constant being given by

$$C = \frac{N\mu^2}{3Vk} \tag{16.27}$$

† The Bohr magneton is the characteristic atomic unit of magnetic dipole moment. In (16.24), e and m denote the electron charge and mass, c is the speed of light, and $\hbar$ is Planck's constant divided by 2π.

The values of χ are generally reported on a molar basis. For 1 mole we take $N = N_0$, Avogadro's number. Then V is the volume of 1 mole, and may be expressed in terms of the mass density ρ and molecular weight M through the relation $\rho = M/V$. The molar value of χ predicted by the Langevin analysis is

$$\chi_{\text{mol}} = \frac{\rho N_0 \mu^2}{3MkT} \tag{16.28}$$

....

16.3.2 Show that the Langevin function

$$\mathscr{L}(x) = \coth(x) - \frac{1}{x}$$

has the limiting forms

(a) $\mathscr{L}(x) \simeq (1/3)x, \qquad x \ll 1,$
(b) $\mathscr{L}(x) \simeq 1 - (1/x), \qquad x \gg 1.$

For part (a) it is recommended that the power series expansions

$$e^{\pm x} = 1 \pm x + \tfrac{1}{2}x^2 \pm \tfrac{1}{6}x^3 + \cdots$$

be employed to find the approximate form of

$$\coth(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}}$$

for small x.

16.3.3 Praseodymium is a rare-earth element of atomic weight 140.9 and density 6.78 gm/cm^3. Its magnetic susceptibility is 5.01×10^{-3} cgs units/mole, at a temperature of 293 K. Determine the magnetic moment of the Praseodymium atom. Express your result in units of the Bohr magneton μ_B.

....

Langevin's classical theory has been superceded by Brillouin's quantum theory of paramagnetism. The physical pictures are much the same for both the classical and quantum theories. Experimentally, very low temperatures and very strong magnetic fields are needed to distinguish between the predictions of the Langevin and Brillouin theories. A strong field and a low temperature cooperate to produce a sizable value of the parameter $\mu H/kT$. When μH becomes comparable to kT, the effects of quantization become discernible and Brillouin's quantum theory of paramagnetism is required to obtain agreement with experiment.

....

16.3.4 Quantum effects become important in paramagnetic materials when the parameter $\mu H/kT$ reaches (or exceeds) values near unity. Assuming that μ is roughly 1 Bohr magneton, what magnetic field strengths are needed to study quantum paramagnetism at

(a) liquid helium temperature (4 K)?
(b) liquid nitrogen temperature (80 K)?
(c) room temperature (300 K)?

Are the required magnetic fields presently attainable?

....

REFERENCES

The thermodynamic and kinetic bases for the Clausius–Clapeyron equation are compared in:

R. P. Feynman, R. B. Leighton, and M. Sands, "The Feynman Lectures on Physics," Volume 1, Lecture 45. Addison-Wesley, Reading, Massachusetts (1963).

The Langevin theory of paramagnetism is developed in:

F. W. Sears, "An Introduction to Thermodynamics, The Kinetic Theory of Gases, and Statistical Mechanics," 2nd ed., Chapter 15. Addison-Wesley, Reading, Massachusetts (1953).

Chapter 17

Quantum Physics and Quantum Statistics

17.1 INTRODUCTION

The microscopic world is the realm of quantum phenomena. At the microscopic level the fundamental properties of matter are found in packages which are always integral multiples of some basic unit. For example, electric charge is *quantized*—it comes in discrete quantities. The smallest unit of electric charge has the value[†]

$$e = 4.803 \times 10^{-10} \text{ esu}$$

This is the charge carried by a proton, for example. A charge of equal magnitude but opposite sign resides on the electron. All distributions of electric charge are collections of these elementary quanta of electric charge.

The discrete nature of mass is also a familiar but perhaps unappreciated aspect of quantization. At the macroscopic level mass distributions appear continuous. On the microscopic level we know that matter is granular—we believe in atoms. Each atom has a characteristic mass, as does each of its component parts, the electrons, protons, and neutrons. So, when we speak of a quantity as being quantized, we mean that it is observed in nature only in integral multiples of some basic unit.

[†] In the mks system, $e = 1.609 \times 10^{-19}$ C.

Modern atomic and nuclear physics have revealed several other quantal features of nature. Of special significance for statistical mechanics is the quantum nature of angular momentum and energy. To develop a feeling for certain aspects of quantum theory we trace its early history in the next section. The student who wishes to recapture a sense of the excitement, the controversy, and the revitalization of physics which followed Max Planck's introduction of the quantum of action will find the articles by Martin J. Klein (1966) and K. K. Darrow (1952) most enlightening.

17.2 THE ORIGINS OF QUANTUM THEORY

Toward the end of the 19th century, several inadequacies of classical physics became apparent. Among these shortcomings was the failure of the equipartition of energy theorem to correctly describe the temperature variations in the heat capacity of diatomic gases. Nor did the specific heats of good electrical conductors follow the equipartition theorem. For some inexplicable reason the mobile electrons were able to transport energy and electric charge without making a detectable contribution to the specific heat. Gibbs' paradox concerning the entropy of mixing was another warning that all was not right with classical statistical physics.

Quantum theory was born, rather hesitantly, with Max Planck's investigation of still another riddle of nature—the blackbody radiation problem. In general, electromagnetic radiation reaching a material surface is partially absorbed and partially reflected. A blackbody is one which completely absorbs any electromagnetic radiation reaching its surface—a perfect absorber. Blackbody radiation may be studied experimentally using an electric furnace whose inner walls can be held at a constant temperature. A small hole in one wall will permit a tiny fraction of the thermal radiation to escape and be sampled. Any light which travels inward through the hole will almost certainly be absorbed before it has a chance to escape, whether or not the furnace walls are black. It is thus the hole in the furnace wall which functions as a blackbody. In principle, a hole in the side of a shoebox, a doorway leading to a darkened room, or any similar cavity opening would serve equally well. Because the radiation trapped in an isothermal cavity is identical with that emitted by a blackbody, the terminology cavity radiation is often used.

Under steady-state conditions, a blackbody emits and absorbs radiant energy at the same rates. The absorbed and emitted radiations may have quite different characteristics. The character of the radiation reaching the blackbody is determined by its sources. It may be monochromatic or a blend of different wavelengths. The flux of radiation emitted by a blackbody depends exclusively on one quantity—the temperature of the blackbody. The rate at which energy is

radiated per unit area (total energy flux) is termed the *surface flux* and is designated by the symbol F. For a blackbody, F is proportional to the fourth power of its absolute temperature

$$F = \sigma T^4 \tag{17.1}$$

This equation is known as the *Stefan–Boltzmann Law*; the constant σ is Stefan's constant

$$\sigma = 5.67 \times 10^{-5} \text{ erg/cm}^2\text{-sec-K}^4 \tag{17.2}$$

Stefan formulated this relation in 1879 on the basis of experimental data. In 1884, Boltzmann used a novel Carnot cycle and classical thermodynamics to prove that F was proportional to T^4, although his analysis could not predict the value of σ.

The spectrum of a blackbody is continuous, that is, all frequencies are present. If a blackbody is scanned by a detector sensitive to a narrow range of frequencies, the surface flux F can be analyzed to give what is called the monochromatic emissive power, $e(\nu)$. The monochromatic emissive power is defined by the statement that $e(\nu)\,d\nu$ is the portion of the flux composed of radiation in the frequency range from ν to $\nu + d\nu$. Integrating $e(\nu)\,d\nu$ over all frequencies gives F

$$F = \int_0^\infty e(\nu)\,d\nu \tag{17.3}$$

In Figure 17.1, $e(\nu)$ is plotted versus ν for several different blackbody temperatures. When Planck began his theoretical work on the blackbody spectrum,

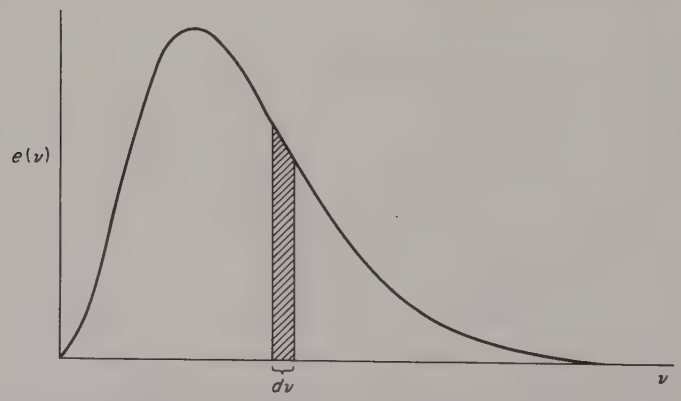


Figure 17.1. The monochromatic emissive power $e(\nu)$ for a blackbody. The hatched area $e(\nu)\,d\nu$ equals the intensity contributed by radiation in the range of frequency from ν to $\nu + d\nu$.

only the high-frequency end of the spectrum had been investigated experimentally. The problem confronting Planck, and others, was to formulate a theory which predicted the shape of the blackbody spectrum—a theory which predicted the functional form of $e(\nu)$.

In 1887 Hertz generated and detected electromagnetic waves, thereby confirming Maxwell's theory of light. Planck envisioned the surfaces of a radiating body as containing innumerable electrical oscillators. Each Hertzian oscillator had a characteristic frequency. An oscillator could absorb radiation of the same frequency it emitted. Thermal equilibrium in a cavity was reached when the rate of emission by each oscillator matched its rate of absorption.

Initially, Planck sought to derive the semiempirical relation for $e(\nu)$ advanced by Wien in 1896. Wien's expression for $e(\nu)$ is

$$e(\nu) = A\nu^3 e^{-B\nu/T} \tag{17.4}$$

where A and B are constants and T is the blackbody temperature. The available experimental data were in close agreement with Wien's formula. Planck assumed that Wien's expression for $e(\nu)$ was the correct one and managed to "derive" it, but only by a process of working backward from the answer—a trick mastered by virtually every student! Early in 1900, experiments by Rubens and Kurlbaum, two of Planck's colleagues, and by Lummer and Pringsheim, extended the measurements of $e(\nu)$ into the low frequency range (infrared). These measurements were in clear disagreement with Wien's formula.

In October, 1900, Planck presented a revised form of $e(\nu)$ which agreed with all data then available—and all accumulated since. Planck had no theory but advanced his formula as the simplest modification of Wien's formula which was in agreement with the data and consistent with thermodynamic considerations. The expression for $e(\nu)$, which proved to be the correct one, and for which Planck as yet had no theory, is

$$e(\nu) = \frac{A\nu^3}{e^{B\nu/T} - 1} \tag{17.5}$$

....

17.2.1 Show that for sufficiently high frequencies Planck's form of $e(\nu)$ reduces to Wien's expression (17.4).

....

Once again convinced that he had the right answer, Planck plunged ahead with the task of deriving this result. To do so he adopted Boltzmann's statistical approach. This was painful for Planck. He had long adhered to the belief that the laws of thermodynamics were of a causal nature—that they could be

derived from the known laws of mechanics and electromagnetism without introducing assumptions of a statistical nature. Boltzmann was able to derive the laws of thermodynamics only by adding a statistical postulate; namely, that the observed properties of a system should correspond to the most probable value, or an appropriate average over a distribution of values. During the course of the analysis Planck found it mathematically advantageous to assign discrete energies to the oscillators. The energies open to oscillators of a given frequency were permitted to be multiples of some basic unit ε. Thus the energy of an oscillator of frequency ν was restricted to be

$$0, \varepsilon, 2\varepsilon, 3\varepsilon, \ldots, n\varepsilon, \ldots, \qquad n = \text{integer}$$

This trick had been employed previously by Boltzmann. Its use reflects the fact that it is conceptually easier to perform addition than it is to integrate. However, when the development reaches the point where mean values must be computed, summations are replaced by integrations, and so on. Our evaluation of the ideal gas partition function in Section 15.8 is an example of such a switch from a discrete to a continuous variable.

To his astonishment, Planck found that he could derive (17.5) only by retaining a finite value for ε. Furthermore, the form of ε was determined. The characteristic energy had to be directly proportional to the oscillator frequency ν. Thus for the first time there was written the now-famous relation

$$\varepsilon = h\nu \tag{17.6}$$

Planck had derived the blackbody spectrum by quantizing energy—by restricting the oscillator energies to multiples of the energy quantum $h\nu$. The constant h is called *Planck's constant* and has the value

$$h = 6.624 \times 10^{-27} \text{ erg-sec}$$

....

17.2.2 Determine the dimensions of h. What other physical quantity has the same dimensions?

....

Incredible as it may seem, Planck initially did not believe in energy quantization! He was convinced that quantization was a mathematical artifice—a neat trick which led to the correct blackbody spectrum. Planck spent the next 10 years retreating from quantization—trying to find an alternate route to the blackbody radiation law.

One of the major reasons for his refusal to embrace energy quantization was its consequences for electromagnetic theory. If the state of an oscillator is quantized, then it can emit and absorb only discrete quanta of energy. This in turn suggests that the energy of the electromagnetic field is quantized. Planck

was not about to scrap the immensely successful theories of Maxwell and Huygens for the sake of the blackbody radiation law.

By 1910, physicists were more or less polarized on the quantization issue. The conservative element, led by Sir James Jeans, maintained that classical physics was "OK" as it stood. Jeans argued that blackbody radiation was not in thermodynamic equilibrium with the radiating surface and so could not be expected to follow the equipartition theorem. The radical elements, championed by Einstein, eagerly accepted the idea of energy quantization. Einstein argued that light quanta were a necessary consequence of Planck's theory. If an oscillator absorbs and emits electromagnetic energy in packets of size $h\nu$, then it would seem that the electromagnetic field is merely a collection of energy parcels—or photons as they are now called. In 1905, Einstein used the photon theory of light to explain the photoelectric effect, an achievement which later earned him the Nobel Prize.

In 1913, Niels Bohr advanced his quantum theory of the hydrogen atom. Both energy and angular momentum are quantized in the Bohr theory. Bohr's theory was subsequently modified and eventually replaced (in 1925–1926) by the less mechanistic modern quantum theory.

Today we have a relativistic quantum theory, thanks to the pioneering work of Dirac and Pauli some forty years ago. Quantum electrodynamics, a quantized electromagnetic theory—inconceivable to Planck—has been with us for 25 years, largely through the efforts of Bethe, Feynmann, Schwinger, and Tomonaga.

Quantum theory has had a profound effect on statistical physics. The following sections sketch the basic modifications wrought by quantum theory.

17.3 ANGULAR MOMENTUM AND STATISTICS

In classical mechanics it proves extremely fruitful to introduce the concept of angular momentum. Its importance stems in part from the fact that the angular momentum of a torque-free system is conserved; it is a constant quantity. For example, the conservation of angular momentum is the basis for Kepler's second law of planetary motion. However, the angular momentum of a planet is composed of two parts. Take Earth as a typical example. Our planet has orbital angular momentum by virtue of its yearly motion around the Sun. In addition it has a spin angular momentum by virtue of its daily rotation. For Earth and the other members of our solar system the spin angular momentum is always small by comparison with the orbital angular momentum.

In his 1913 theory, Bohr envisioned the hydrogen atom as a miniature solar system with the proton serving as a sun and the electron revolving about it in

a planetary orbit. A basic postulate of Bohr's theory was that the orbital angular momentum of the electron was quantized, that is, it could take on only certain discrete values given by

$$L_{\text{orb}} = n\hbar, \qquad n = 1, 2, 3, \ldots$$

The quantity $\hbar$ (read "h cross") is Planck's constant divided by 2π

$$\hbar = \frac{h}{2\pi} = 1.054 \times 10^{-27} \text{ erg-sec} \tag{17.7}$$

The Bohr theory did not survive—it was replaced by a more satisfactory quantum theory. However, the quantization of orbital angular momentum forms a key part of the newer theory as well.

In 1925 Uhlenbeck and Goudsmit postulated a quantum mechanical analog of the planetary spin angular momentum. They noted that certain discrepancies between theoretical and experimental spectroscopy could be removed by assigning to the electron an intrinsic angular momentum over and above any possible orbital angular momentum. In particular, they showed that theory and experiment would be brought into accord if the electron had an intrinsic angular momentum of $\frac{1}{2}\hbar$. The terminology spin angular momentum was adopted to describe this new property because one could envision an electron not only moving about the atomic nucleus in a planetary-like orbit but also spinning rapidly about an axis of its own as does the Earth. The picture of an electron spinning rapidly about its own axis is a picturesque analogy, but a misleading one. The angular speed of rotation required to produce an angular momentum of $\frac{1}{2}\hbar$ is so large that it would give the surface of the electron a speed in excess of the speed of light. So, electrons have a spin angular momentum just because they are electrons, not because they are spinning spheres. Nevertheless, the name spin angular momentum has persisted. Uhlenbeck and Goudsmit introduced the spin angular momentum of the electron in a purely *ad hoc* manner. They offered no theoretical basis for the existence of the electron spin momentum. It remained for P. A. M. Dirac to show, in 1928, that the spin of $\frac{1}{2}\hbar$ emerged automatically by requiring that the quantum theory of the electron be in accord with the principles of the special theory of relativity.

The other two primary constituents of matter, the neutron and proton, also have an intrinsic (spin) angular momentum of $\frac{1}{2}\hbar$. The electron, proton, and neutron are often referred to as "spin $\frac{1}{2}$" particles.

Particles with spin other than $\frac{1}{2}$ exist. For example, the α particle (the nucleus of a helium atom) is a spin zero particle; it has no intrinsic spin angular momentum. The pion (pi meson) is another spin zero particle. The photon is a spin 1 particle; it possesses an intrinsic angular momentum of $\hbar$. Gravitons

are spin 2 particles.† The spins of many different atomic nuclei have been measured. Spin values of 0, $\frac{1}{2}$, 1, $\frac{3}{2}$, 2, $\frac{5}{2}$, and so on, have been detected.

The quantum unit of spin angular momentum is $\frac{1}{2}\hbar$. The spin angular momentum of every particle can be expressed as an integral multiple of $\frac{1}{2}\hbar$ (including zero). Nature has seen fit to distinguish rather sharply between particles of integral spin and those of odd half-integral spin. Entities with integral spin are termed *bosons*; those with odd half-integral spin are called *fermions*. Thus, electrons, protons, and neutrons are fermions while photons, gravitons, and α particles are bosons.

It would carry us too far afield to delve further into the foundations of the fermion–boson scheme of classifying particles. It suffices to state that one set of quantum mechanical rules applies to a system of fermions and a different set is applicable to a collection of bosons. There is a carryover to statistical physics—there are two versions of quantum statistical mechanics. The so-called Bose–Einstein statistics version applies to systems composed of bosons. For a collection of fermions the appropriate statistical mechanics is called Fermi–Dirac statistics. Fermi–Dirac statistics is developed in the next section. Bose–Einstein statistics is treated in Section 17.5.

The word "statistics" which follows Bose–Einstein and Fermi–Dirac is a truncation of statistical mechanics. It is not meant to imply that different statistical methods are involved. Rather, two different quantum mechanical constraints lead to different particle distributions via the same statistical arguments.

17.4 FERMI–DIRAC STATISTICS

Consider a system of N identical fermions, and mentally chop its phase space into a great many cells of equal volume. With each cell there is associated a definite energy—the energy of a particle in that cell. As we noted in Chapter 15, an important assumption is implicit in assigning a definite energy to each segment of phase space. Assigning a particle a certain energy because it is in a particular region of phase space ignores the effects of other particles and thereby restricts us to situations in which mutual interactions can be ignored. We deal with an independent-particle model. Even when mutual interactions are not ignorable the present approach can be generalized in a fruitful fashion.

---

† The quanta associated with gravitation are termed *gravitons*. The weakness of gravity has prevented their detection as yet, but theory dictates that their spin angular momentum be $2\hbar$.

The Boltzmann artifice of treating the energies as discrete becomes a necessity in QSM[†]—cell energies are quantized. In many instances the energy spectrum is very compact; there are many phase-space cells whose energies differ only slightly. We therefore consider groups of cells on a wholesale basis and assign the same energy to each. In QSM the distinction between fermions and bosons arises with respect to the capacity of the phase-space cells. Each phase-space cell has a maximum capacity of one fermion. Fermions are terribly introverted—they demand privacy.[‡] Thus if we look at the phase space of a fermion system we see some cells with one occupant while the remaining cells are empty. The introverted nature of fermions is a manifestation of the *Pauli exclusion principle* which decrees that identical fermions cannot occupy the same quantum state. To define precisely what is meant by a quantum state would take us far afield of our purpose here. In effect each phase-space cell corresponds to a distinct quantum state. In this connection, we remark that quantum physics forces an enlargement of the number of variables needed to specify the address of a phase-space cell. In addition to the position and linear momentum variables we must specify the value of a spin variable. Geometrically the spin angular momentum variable is represented as a vector, with different spin states corresponding to different orientations of the vector. For a particle of spin S there are $2S+1$ permitted orientations, that is, $2S+1$ distinct spin states. For electrons and other spin $\frac{1}{2}$ particles there are two possible spin states. The two permitted orientations of the spin angular momentum vector are antiparallel and the states are usually referred to as the spin up and spin down states.

If fermions are introverts, bosons are extroverts for there is no limit to the number of bosons which a given cell can accommodate. In a boson system some phase-space cells might be empty, others might contain a single boson. Still others might house two, three, and so on. It is a peculiarity of Bose–Einstein statistics that each phase-space cell must be prepared to accommodate even an infinite number of boson guests.

There is one other essential difference between CSM and QSM. In developing CSM we carefully numbered the atoms. We took great care to count the number of different ways of arranging the atoms. Alas, atoms do not obey the laws of classical mechanics. They do not, like pool balls, bear different numbers and

---

[†] Henceforth we use QSM to signify quantum statistical mechanics and CSM to stand for classical statistical mechanics.

[‡] We *do not* wish to suggest that the fermions have a will or that they are conscious. Such a suggestion was advanced recently. See A. A. Cochran, Relationships Between Quantum Physics and Biology, *Foundations Phys.* **1**, 235 (1971). Needless to say, Cochran's views did not go unchallenged. See J. F. Woodward *et al.*, Photon Consciousness: Fact or Fancy? *Foundations Phys.* **2**, 241 (1972).

colors. Atoms of the same element are indistinguishable.† QSM formally recognizes the meaning of identical particles. In CSM the microstates depicted in Figure 17.2 are distinct. In QSM they are identical. There is no way to distinguish the two atoms, hence we must remove the numbers. With the numbers removed, the two microstates of Figure 17.2 are identical. So, when it comes time to count microstates in QSM we must take care not to indulge in double-entry bookkeeping. We must count only distinct microstates.

Figure 17.2. Two microstates: With the atoms numbered the microstates are distinguishable. Removing the numbers from the atoms renders the microstates identical.

Consider the ith group of phase-space cells, Q_i in number. Each of these cells is characterized by the same energy ε_i. Let N_i denote the number of fermions distributed over the Q_i cells. The number of distinct microstates associated with a particular distribution of fermions is given by

$$W_i = \frac{Q_i!}{N_i!(Q_i - N_i)!} \tag{17.8}$$

To trace the origin of this assertion we imagine ourselves building up the W_i microstates by tossing fermions into phase-space cells one at a time. We first pretend that the fermions are distinguishable, that is, that they bear numbers $1, 2, 3, \ldots, N_i$. We can correct for the indistinguishability as a final step. We start with Q_i empty cells. The first fermion can be placed in any of the Q_i cells. The second fermion can be placed in any of the remaining $Q_i - 1$ empty cells. The one occupied cell is off limits for a second fermion—remember, they are introverts. Collectively there are $Q_i(Q_i - 1)$ ways of distributing the first two fermions. For each of these $Q_i(Q_i - 1)$ ways a third fermion can take its pick of the $Q_i - 2$ empty cells. Thus there are $Q_i(Q_i - 1)(Q_i - 2)$ ways of distributing three fermions. Pushing on, the N_ith fermion finds only $(Q_i - N_i + 1)$ empty cells from which to choose. The total number of ways of distributing the N_i fermions, assuming them to be distinguishable, is

$$Q_i(Q_i - 1)(Q_i - 2)\cdots(Q_i - N_i + 1) = \frac{Q_i!}{(Q_i - N_i)!}$$

....

17.4.1 Establish the above equality.

....

† This presumes that they are not different isotopes of the element.

We correct for the indistinguishability of the fermions by dividing by $N_i!$ That is, there are $N_i!$ rearrangements of N_i distinguishable particles, all of which become identical when the particles are rendered indistinguishable. Thus the $Q_i!/(Q_i-N_i)!$ ways enumerated above counts each distinct distribution $N_i!$ times, whence,

$$W_i = \frac{Q_i!}{N_i!(Q_i-N_i)!} \tag{17.8}$$

is the number of microstates corresponding to a distribution of N_i fermions over Q_i cells. If there are M groups of cells, the grand total of possible microstates is

$$W\{N_i\} = \prod_{i=1}^{M} W_i = \prod_{i=1}^{M} \frac{Q_i!}{N_i(Q_i-N_i)!} \tag{17.9}$$

....

17.4.2 Explain the logic which dictates that $W = \prod_{i=1}^{M} W_i$. Why is the total number of microstates not merely the sum of the W_i?

....

The procedure for finding the most probable distribution follows the lines established in CSM. We maximize $\ln W$ subject to the constraints that

$$\sum_{i=1}^{M} N_i = N \tag{17.10}$$

and

$$\sum_{i=1}^{M} N_i \varepsilon_i = U \tag{17.11}$$

To this end we introduce the most probable distribution

$$\{N_i\} = N_1, N_2, \ldots, N_k, \ldots, N_s, \ldots, N_M$$

and an adjacent distribution

$$\{N_i'\} = N_1, N_2, \ldots, N_k - 1, \ldots, N_s + 1, \ldots, N_M$$

This automatically guarantees that both distributions correspond to the same total number of particles. The most probable distribution follows by requiring that

$$\ln W\{N\} = \ln W\{N_i'\} \tag{17.12}$$

and that, simultaneously,

$$U = \sum_{i=1}^{M} N_i \varepsilon_i = \sum_{i=1}^{M} N_i' \varepsilon_i$$

As in CSM we can write

$$\ln W\{N_i\} - \beta \sum_{i=1}^{M} N_i \varepsilon_i = \ln W\{N_i'\} - \beta \sum_{i=1}^{M} N_i' \varepsilon_i$$

$\ln W'$ is given by a similar expression. Inserting the explicit forms for W_i and W_i' gives

$$\sum_{i=1}^{M} \ln \frac{Q_i!}{N_i!(Q_i - N_i)!} - \beta N_i \varepsilon_i = \sum_{i=1}^{M} \ln \frac{Q_i!}{N_i'!(Q_i - N_i')!} - \beta N_i' \varepsilon_i$$

As only the kth and sth groups of cells differ in composition there is a wholesale cancellation of terms on either side, leaving

$$\ln\left(\frac{Q_k - N_k + 1}{N_k}\right) - \beta\varepsilon_k = \ln\left(\frac{Q_s - N_s}{N_s + 1}\right) - \beta\varepsilon_s$$

It is legitimate to replace $Q_k - N_k + 1$ by $Q_k - N_k$ and $N_s + 1$ by N_s, leaving

$$\ln\left(\frac{Q_k}{N_k} - 1\right) - \beta\varepsilon_k = \ln\left(\frac{Q_s}{N_s} - 1\right) - \beta\varepsilon_s$$

However, k and s are arbitrary so that $\ln((Q_i/N_i) - 1) - \beta\varepsilon_i$ must have a common value for all M groups of cells. We denote this value as $-\beta\mu$, thereby obtaining

$$\ln\left(\frac{Q_i}{N_i} - 1\right) - \beta\varepsilon_i = -\beta\mu, \qquad i = 1, 2, \ldots, M$$

Solving for N_i results in

$$N_i = \frac{Q_i}{\exp[\beta(\varepsilon_i - \mu)] + 1} \tag{17.13}$$

This is the sought-for Fermi–Dirac distribution. The quantity μ is the chemical potential while β proves to be $1/kT$ as before. Mathematically β and μ enter as quantities which guarantee satisfaction of the conditions

$$N = \sum N_i$$

and

$$U = \sum N_i \varepsilon_i$$

The basic relation (17.13) allows a most fruitful interpretation. The quantity

$$\frac{N_i}{Q_i} = \frac{1}{\exp[\beta(\varepsilon_i - \mu)] + 1} \tag{17.13a}$$

is the fraction of the Q_i cells of energy ε_i which are occupied. Statistically we

can interpret N_i/Q_i as the probability that any particular cell of energy ε_i is occupied by a fermion. In many situations the index i need not be displayed explicitly. We shall refer to

$$\frac{N}{Q} = \frac{1}{e^{\beta(\varepsilon-\mu)}+1} \tag{17.14}$$

as the *occupation probability*. In situations where the discrete nature of the energy spectrum is of no consequence, we can use the technique of converting sums to integrals and write the occupation probability as

$$\frac{dN}{dQ} = \frac{1}{e^{\beta(\varepsilon-\mu)}+1} \tag{17.14a}$$

where dN is the expected number of fermions in the dQ phase-space cells characterized by energy ε.

One of the earliest applications of Fermi–Dirac statistics was made by Arnold Sommerfeld in 1928 to the problem of determining the heat capacity of electrical conductors. A conducting solid is viewed as a collection of three-dimensional simple harmonic oscillators, with each atom possessing six degrees of freedom. The empirical law of Dulong–Petit states that the molar heat capacity of a solid is 6 cal/mole-K. This can be understood on the basis of energy equipartition. The equipartition theorem assigns a molar heat capacity of $\frac{1}{2}R$ to each degree of freedom. With six degrees of freedom for the three-dimensional simple harmonic oscillator (see Figure 11.4)

$$C_v = 3R \simeq 6 \text{ cal/mole-K}$$

The paradox arises because many of the solids which have $C_v \simeq 6$ cal/mole-K are also good electrical conductors (Ag, Au, Cu, Na, and so on). To explain their high electrical conductivity it is necessary to hypothesize that many of the atoms forming the crystalline lattice have sacrificed an electron. It is the directed motions of these electrons which constitute an electric current. Early models could explain the observed electrical conductivity by assuming that virtually every atom contributed one electron. Such an electron gas should also contribute to the heat capacity. Heat absorbed by the solid should be shared by the electrons and by the lattice of ions and atoms. It was assumed that the electrons behaved like a monatomic gas. The equipartition theorem predicted that they should have a molar heat capacity of $\frac{3}{2}R$. Virtually every atom donates an electron, leaving an ionic lattice. Each lattice ion (or atom) contributes six degrees of freedom. With each atom contributing one electron the total molar heat capacity should be $\frac{3}{2}R$ for the electrons plus $3R$ for the lattice, giving

$$C_v = \tfrac{9}{2}R$$

Experimentally, however, $C_v \simeq 3R$; the electrons apparently do not "play ball." Their contribution of $\frac{3}{2}R$ to the heat capacity is missing. Yet, the electrons seem to behave normally insofar as the electrical conductivity is concerned.

The apparent paradox evaporates once Fermi–Dirac statistics are brought to bear, and for this reason: Electrons distribute themselves over phase space according to the Fermi–Dirac expression (17.14), and not according to the classical Boltzmann distribution law (15.31). As a consequence of their Fermi–Dirac distribution, only a tiny fraction of the electrons are free to increase their energy and thereby share in the partitioning of energy. The electron degrees of freedom are suppressed and the ionic lattice receives virtually all of the absorbed thermal energy. What prevents the electron gas from absorbing thermal energy? Briefly, it is the Pauli exclusion principle. The exclusion principle forbids electrons from occupying the same phase-space cell. Since most of the electrons are unable to find a vacant phase-space cell of higher energy they are unable to partake of the absorbed energy. The small fraction that find empty phase-space cells nearby are able to absorb energy. Because they are relatively few in number their contribution to the heat capacity of the solid is ordinarily much smaller than that made by the lattice. We must study the Fermi–Dirac distribution in depth to see just how the exclusion principle operates to insulate the electron gas.

For openers, let us estimate the value of the chemical potential μ appropriate to the electron gas in a good electrical conductor. As it turns out, a typical value for μ is a few electron volts. This is extremely large by comparison to kT, which is about 1/40 eV at room temperature. In general, the value of μ depends on the concentration of electrons and the temperature. As a very good first approximation we can take $T = 0$ for the purposes of estimating μ. It has become customary to refer to the value of μ at absolute zero as the Fermi energy ε_F

$$\mu(T=0) \equiv \varepsilon_F$$

In many instances it is permissible to use ε_F in place of μ even at room temperature because the chemical potential varies only weakly with the temperature.

With $T = 0$, the parameter $\beta = 1/kT$ is infinite and

$$\exp[\beta(\varepsilon-\varepsilon_F)] + 1 = \begin{cases} +1, & \varepsilon < \varepsilon_F \\ \infty, & \varepsilon > \varepsilon_F \end{cases}$$

The occupation probability is given by

$$\frac{N}{Q} = \frac{1}{\exp[\beta(\varepsilon-\varepsilon_F)] + 1} = \begin{cases} 1, & \varepsilon < \varepsilon_F \\ 0, & \varepsilon > \varepsilon_F \end{cases}$$

Figure 17.3 is a plot of N/Q versus ε at $T = 0$. The figure simply reveals that electrons fill the cells of lowest energy consistent with the Pauli exclusion principle.

The value of the Fermi energy is determined by the condition that the total number of electrons equal the total number of phase-space cells whose energy does not exceed ε_F. In QSM it is preferable to deal with a phase space which combines position and momentum rather than position and velocity. Because the Fermi distribution is isotropic in momentum space we employ spherical momentum space variables. Cells corresponding to momenta in the spherical shell defined by p and $p+dp$ occupy a phase-space volume of

$$8\pi p^2\, dp\, V$$

A factor of 2 enters because of the electron spin. Thus for the momentum p to $p+dp$ there is a phase-space volume $4\pi p^2\, dp\, V$ open to electrons with spin up and an equal volume open to electrons with spin down. The volume of the unit cell is h^3, where h is Planck's constant. Dividing the phase-space volume by the volume per cell gives the number of cells

$$dQ = \frac{8\pi p^2\, dp\, V}{h^3} \tag{17.15}$$

The electron energy and momentum are related by

$$\varepsilon = \frac{p^2}{2m}$$

so that the radius of the momentum-space sphere corresponding to the energy ε_F is given by

$$p_F = \sqrt{2m\varepsilon_F}$$

The quantity p_F is called the *Fermi momentum*. The surface of the momentum-

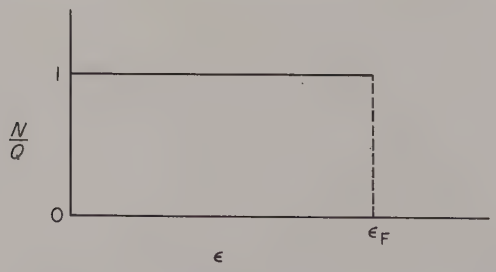


Figure 17.3. The occupation probability N/Q at absolute zero. Phase-space cells of energy greater than the Fermi energy ε_F are empty; those of energy up to ε_F are completely filled.

space sphere of radius p_F is referred to as the *Fermi surface*. At absolute zero, all electrons are beneath the Fermi surface. The total number of occupied cells is

$$Q = \frac{8\pi V}{h^3}\int_0^{p_F} p^2\,dp = \frac{8\pi V}{3h^3}(2m\varepsilon_F)^{3/2}$$

Since the N electrons completely fill these cells we set $Q = N$ and obtain for the Fermi energy

$$\varepsilon_F = \frac{h^2}{2m}\left(\frac{3n}{8\pi}\right)^{2/3} \tag{17.16}$$

where m is the electron mass and $n = N/V$ is the electron concentration (number density). For a good conductor, n is likely to exceed 10^{22} electrons/cm^3, which places ε_F in the electron volt range. For future reference, we remark that the Fermi momentum is independent of the mass of the particle. The value of p_F is fixed by the number density n. A convenient expression for p_F is

$$p_F = \hbar(3\pi^2 n)^{1/3}, \qquad \hbar = \frac{h}{2\pi} \tag{17.17}$$

....

17.4.3 The density of metallic sodium is 0.971 gm/cm^3 and its atomic weight is 23.0. Assume that each sodium atom donates one electron and compute the Fermi energy of the electron gas.

....

Example 17.1 Average Energy versus Fermi Energy

Even if the electron gas is at absolute zero, the mean kinetic energy of the electrons is quite large—another consequence of the Pauli exclusion principle. To show this, we compute the mean kinetic energy of the electron gas at $T = 0$.

All phase-space cells are occupied for $\varepsilon < \varepsilon_F$ and all are empty for $\varepsilon > \varepsilon_F$. We can therefore write

$$\langle\varepsilon\rangle = \frac{1}{N}\int_{\substack{\text{all}\\ \text{cells}}} \varepsilon\,dN$$

where

$$dN = \begin{cases} dQ, & \varepsilon < \varepsilon_F \\ 0, & \varepsilon > \varepsilon_F \end{cases}$$

dQ is the number of phase-space cells for energies in the range ε to $\varepsilon + d\varepsilon$. If we write

$$dQ = \frac{8\pi}{h^3}Vp^2\,dp$$

and then use $\varepsilon = p^2/2m$ and its differential $d\varepsilon = p\,dp/m$, we find

$$dQ = \frac{4\pi V}{h^3}(2m)^{3/2}\sqrt{\varepsilon}\,d\varepsilon$$

Observing that $N = \int dN = \int_0^{\varepsilon_F} dQ$ we obtain

$$\langle\varepsilon\rangle = \frac{\int_0^{\varepsilon_F} \varepsilon^{3/2}\,d\varepsilon}{\int_0^{\varepsilon_F} \varepsilon^{1/2}\,d\varepsilon} = \frac{3}{5}\varepsilon_F$$

For good electrically conducting solids ε_F is typically a few electron volts. The corresponding electron speed is about 1000 times the speed of sound. The exclusion principle really keeps them jumping—even at absolute zero! ✧

....

17.4.4 (a) Show that the volume of a spherical shell in momentum space may be expressed in terms of the energy ε by

$$4\pi p^2\,dp = 4\pi\sqrt{2}\,m^{3/2}\sqrt{\varepsilon}\,d\varepsilon$$

provided $\varepsilon = p^2/2m$.

(b) Supply reasoning to establish that the equation relating μ to the number density n and temperature is

$$n = \frac{4\pi(2m)^{3/2}}{h^3}\int_0^\infty \frac{\sqrt{\varepsilon}\,d\varepsilon}{e^{\beta(\varepsilon-\mu)}+1}$$

(c) Show that (17.16) results as a special case of the above equation when $\beta = \infty$ $(T = 0)$.

....

Plots of the occupation probability N/Q versus ε for several values of ε_F/kT are shown in Figure 17.4. Where ε differs from ε_F by more than a few times kT, the exponent $\beta(\varepsilon-\varepsilon_F)$ has a large magnitude and the exponential is either very large or very small by comparison with unity, that is,

$$\exp[\beta(\varepsilon-\varepsilon_F)]\begin{cases}\simeq \infty, & \varepsilon-\varepsilon_F \gg kT \\ = 1, & \varepsilon = \varepsilon_F \\ \simeq 0, & \varepsilon_F-\varepsilon \gg kT\end{cases} \qquad \beta = \frac{1}{kT}$$

Thus for energies much below ε_F, N/Q is nearly one; N/Q drops sharply to $\frac{1}{2}$ at $\varepsilon = \varepsilon_F$ and tails off exponentially to zero when ε exceeds ε_F. We can describe the situation at $T > 0$ as one in which a small fraction of the electrons have been excited into phase-space cells above the Fermi surface. Most of the

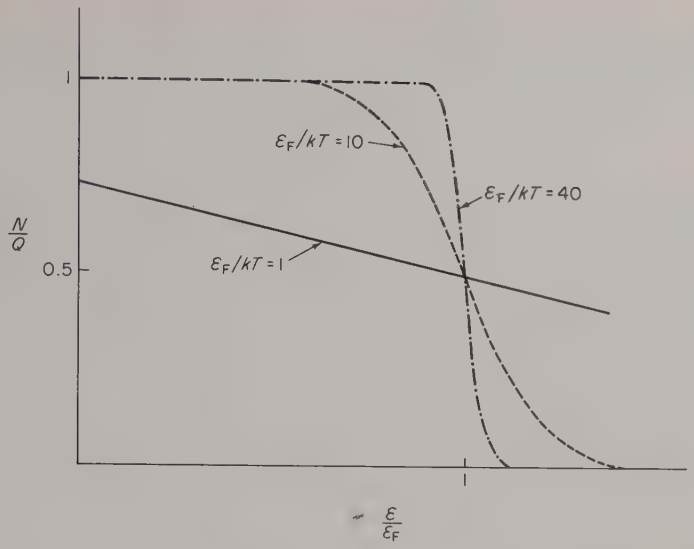


Figure 17.4. The Fermi–Dirac occupation probability for several values of ε_F/kT.

electrons are buried beneath the Fermi surface. They are unable to share in energy absorbed by the material because there is no phase space available to them—no vacant phase-space cells above them on the energy scale. Only the small fraction of the electrons in the vicinity of the Fermi surface can respond to external forces. Thus the electron gas in a metal is more or less insulated from the rest of the solid. With this picture in mind we can estimate the molar heat capacity C_v of the electron gas, which is considerably smaller than the value of (3/2) R predicted by the equipartition theorem.

It is not difficult to write down an exact expression for C_v. Evaluating the resulting integral is a bit tricky. Instead of seeking the exact result, let us determine C_v approximately. Electrons near the Fermi surface behave classically, that is, they move about freely like the atoms of an ideal gas. The molar heat capacity of the electron gas (e.g.) should then be (3/2) R times the fraction of the electrons near the Fermi surface

$$(C_v)_{\text{e.g.}} = \tfrac{3}{2}R(\text{fraction of the electrons near the Fermi surface}) \quad (17.18)$$

A span of energies of order kT above and below ε_F defines the region where there are electrons *and* empty phase-space cells. For $\varepsilon \gg \varepsilon_F + kT$ there are empty cells but virtually no electrons. For $\varepsilon \ll \varepsilon_F - kT$ there are plenty of electrons but no vacant cells.

Occupied energy states span an energy range of roughly ε_F, so that the range of $2kT$ over which vacancies occur encompasses a fraction of order $2kT/\varepsilon_F$ of

all electrons. Taking $2kT/\varepsilon_F$ for the fraction of electrons near the Fermi surface in (17.18) gives

$$(C_v)_{e.g.} \simeq 3\left(\frac{kT}{\varepsilon_F}\right)R \ll \frac{3}{2}R \tag{17.19}$$

A sophisticated calculation using the exact but unwieldy form of N/Q gives a value of $\pi^2/2 \simeq 4.93$ in place of the 3 in (17.19). This result for the electron gas heat capacity resolves the apparent paradox. At ordinary temperatures the electron gas contribution to C_v is undetectably small by comparison to that of the lattice. The electron gas contribution to the heat capacity of metals has been measured in experiments conducted at extremely low temperatures ($\sim$1 K). At low temperatures quantum effects force the lattice heat capacity toward zero. At very low temperatures, the lattice heat capacity is proportional to T^3. As T approaches zero, the electron gas heat capacity—which by (17.19) decreases only linearly with T—dominates the lattice contribution.

Further applications of the Fermi–Dirac statistics are presented in the next chapter.

....

17.4.5 (a) Show that a two-term Taylor series for the occupation probability about $\varepsilon = \varepsilon_F$ gives

$$\{\exp[\beta(\varepsilon-\varepsilon_F)]+1\}^{-1} = \tfrac{1}{2} - \tfrac{1}{4}\beta(\varepsilon-\varepsilon_F)$$

Verify that the right side equals 1 at $\varepsilon = \varepsilon_F - 2kT$ and 0 at $\varepsilon = \varepsilon_F + 2kT$.

(b) What is the average occupation probability over the range of energies from $\varepsilon_F - 2kT$ to $\varepsilon_F + 2kT$?

(c) Can you reconcile the results of (a) and (b) with the text statement that the fraction of electrons near the Fermi surface is $2kT/\varepsilon_F$?

....

17.5 BOSE–EINSTEIN STATISTICS

The phase-space distribution of bosons is determined by precisely the same methods used for fermions. The major difference is that each phase-space cell can accommodate any number of bosons. There is no exclusion principle for bosons.

Consider N_i bosons distributed over Q_i phase-space cells. The Q_i cells correspond to the same energy ε_i. Regard the bosons as distinguishable at first, then correct at the end of the game by dividing by $N_i!$. To identify the Q_i cells we use $Q_i - 1$ walls. To count the number of ways W_i of distributing the N_i

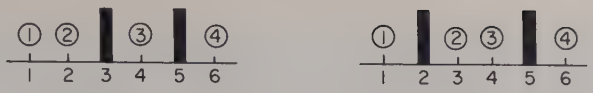


Figure 17.5. Two ways of distributing 4 bosons over 3 cells. The bosons are supposed distinguishable at this stage and so are numbered.

bosons over the Q_i cells we imagine N_i+Q_i-1 *positions* along a line. We place either a boson or a wall at each position. Figure 17.5 illustrates two ways of distributing $N_i=4$ bosons over $Q_i=3$ cells. There are $N_i+Q_i-1=6$ positions; four occupied by bosons, two by cell walls.

Start dropping in bosons—one at a time. When all N_i bosons have been positioned, the remaining Q_i-1 positions are called *walls*. The first boson can be placed at any one of the N_i+Q_i-1 positions, the second at any one of the remaining N_i+Q_i-2, and so on. The N_ith boson can be placed at any of the Q_i positions, leaving Q_i-1 positions for cell walls. Because we regard the particles as distinguishable at this stage, the total number of ways of distributing them is the product

$$(N_i+Q_i-1)(N_i+Q_i-2)(N_i+Q_i-3)\cdots(Q_i)=\frac{(N_i+Q_i-1)!}{(Q_i-1)!}$$

To correct for the indistinguishability we divide by $N_i!$ giving

$$W_i=\frac{(N_i+Q_i-1)!}{N_i!(Q_i-1)!} \tag{17.20}$$

The total thermodynamic probability W is the product of the W_i for all groups of phase-space cells

$$W=\prod_{i=1}^{M} W_i \tag{17.9'}$$

It is left as an exercise for the student (see Problem 17.5.1) to show that the most probable distribution has the form

$$N_i=\frac{Q_i}{\exp[\beta(\varepsilon_i-\mu)]-1} \tag{17.21}$$

This is the Bose–Einstein distribution. By an extension of the arguments presented in Appendix B it can be shown that $\beta=1/kT$. As before, μ is the chemical potential; its value depends on the nature of the bosons, their number density, and the temperature.

In the next chapter we apply the Bose–Einstein statistics to the problem of blackbody radiation and indicate how the theory has recently been used to support arguments for the "big bang" theory of the universe.

....

17.5.1 Use (17.20) in (17.9′) and follow the procedures used with Fermi–Dirac statistics to derive the Bose–Einstein distribution (17.21). The equations of constraint are (17.10) and (17.11).

....

17.6 QSM OR CSM?

The Bose–Einstein and Fermi–Dirac distributions differ in form only in that there is a -1 in the denominator of (17.21) and a $+1$ in the Fermi–Dirac denominator (17.13). Under circumstances where

$$e^{-\beta\mu} \gg 1 \tag{17.22}$$

both Fermi–Dirac and Bose–Einstein statistics emulate the classical Boltzmann statistics. Thus to the extent that the ± 1 is ignorable by comparison with the exponential, N_i is given by the Boltzmann distribution of CSM

$$N_i = \frac{Q_i}{\exp[\beta(\varepsilon_i - \mu)] \pm 1} \simeq \exp(\beta\mu)\, Q_i \exp(-\beta\varepsilon_i) \tag{17.23}$$

The right side is identical with the form of N_i given in (15.33), wherein $N/Z(\beta)$ is the quantity designated as $e^{\beta\mu}$ in (17.23).

At ordinary temperatures and densities (17.22) holds for most systems. Quantum statistical effects are then not in evidence and one can say that QSM reduces to CSM. Quantum effects become evident at low temperatures and/or high densities—conditions which tend to make $e^{-\beta\mu} < 1$.

When $e^{-\beta\mu} \ll 1$, that is, when $\mu \gg kT$, the distribution is said to be degenerate. As the ratio μ/kT increases the distribution degenerates, in this sense; all phase-space cell populations approach values appropriate to the distribution at absolute zero. The metallic electron gas studied in Section 17.4 is degenerate, its distribution differing from the $T = 0$ form only in the vicinity of the Fermi surface.

....

17.6.1 Explain why $\mu/kT \gg 1$ is equivalent to the condition $e^{-\beta\mu} \ll 1$.

....

To determine whether or not a distribution is degenerate one must evaluate the chemical potential μ and then compute $e^{-\beta\mu}$. In general the equation determining the ratio μ/kT is transcendental (compare Problem 17.4.4). Fortunately, it is often not necessary to know the precise value of $e^{-\beta\mu}$, it being sufficient to know whether it is large or small in comparison with unity. A

"quick-and-dirty" test uses (17.23) to evaluate $e^{-\beta\mu}$. That is, one assumes that the Boltzmann distribution gives a satisfactory description and then checks for consistency. Failure to meet the test warns of quantum degeneracy. We first evaluate the nondegenerate expression for $e^{-\beta\mu}$ and then illustrate the technique. From (17.23) the total number of particles follows as

$$N = \sum_i N_i = \exp(\beta\mu) \sum_i Q_i \exp(-\beta\varepsilon_i)$$

The summation is converted to an integration over energy

$$\sum_i Q_i \exp(-\beta\varepsilon_i) = \int dQ \exp(-\beta\varepsilon)$$

Taking $dQ = 4\pi V p^2\, dp/h^3$ and $\varepsilon = p^2/2m$ results in

$$e^{-\beta\mu} = \left(\frac{2\pi mk}{h^2}\right)^{3/2} \frac{T^{3/2}}{N/V} \tag{17.24}$$

This form of $e^{-\beta\mu}$ points up our earlier remark that low temperatures and high densities tend to make $e^{-\beta\mu} \ll 1$ and promote a quantum degeneracy. The result (17.24) can be expressed in a form inviting an intuitive and easily remembered interpretation, namely,

$$e^{-\beta\mu} = \frac{(2\pi mkT)^{3/2} V/N}{h^3} \tag{17.25}$$

The quantity $(2\pi mkT)^{1/2}$ has the dimensions of momentum and is very nearly equal to the rms momentum. The quantity $(2\pi mkT)^{3/2}$ is the characteristic momentum space volume per particle. The quantity V/N is the spatial volume per particle. Their product, the numerator of (17.25), is the representative phase-space volume per particle. The factor h^3 in the denominator of (17.25) is the volume of one phase-space cell. It follows from (17.25) that the quantity $e^{-\beta\mu}$ measures the average number of phase-space cells allotted to each particle. The condition validating the use of classical statistics is $e^{-\beta\mu} \gg 1$, or

$$\frac{\text{volume of phase space per particle}}{\text{volume of one phase-space cell}}$$

$$= \text{number of phase-space cells per particle} \gg 1$$

Classical statistics therefore apply under conditions which produce a sparse population of phase-space cells. Each particle has—on the average—a large number of phase-space cells to roam about in. Conversely we expect quantum degeneracy to appear when the population per phase-space cell approaches unity. We next present examples illustrating this point.

Example 17.2 Degenerate or Nondegenerate?

We consider in turn the cases of (a) nitrogen molecules at normal atmospheric conditions, (b) liquid helium at the normal boiling point, and (c) electrons under typical laboratory plasma conditions. For each of the three we evaluate

$$e^{-\beta\mu} = \frac{(2\pi mkT)^{3/2} V/N}{h^3}$$

$e^{-\beta\mu} \gg 1$ implies that classical statistics is adequate, while $e^{-\beta\mu} \lesssim 1$ warns that quantum degeneracy may invalidate considerations based on CSM.

(a) For N_2 molecules at 1 atm and $T \simeq 300$ K we have (in cgs units)

$$m \simeq 5 \times 10^{-23}, \qquad \frac{V}{N} \simeq 4 \times 10^{-20}, \qquad (2\pi mkT)^{3/2} \simeq 6 \times 10^{-53}$$

$$h^3 \simeq 2 \times 10^{-79}, \qquad e^{-\beta\mu} \simeq 10^7$$

It follows that the air we breathe is adequately described by CSM.

(b) For the atoms of liquid helium at the normal boiling point we have $m \simeq 6 \times 10^{-24}$ and $T \simeq 4.2$ K. The volume per atom may be determined from the mass density of 0.12 gm/cm^3,

$$\frac{V}{N} = \frac{\text{mass per atom}}{\text{mass per unit volume}} \simeq 5 \times 10^{-23}$$

These figures give $e^{-\beta\mu} \simeq 1$. It is therefore not surprising to find that liquid helium exhibits quantum degeneracy effects near its normal boiling point. Indeed, helium exhibits superfluid properties below a temperature of 2.17 K.

(c) As a final example, consider electrons under conditions typical of a laboratory plasma; $m \simeq 10^{-27}$, $N/V \simeq 10^{12}$, $T \simeq 10^3$ K. One finds $e^{-\beta\mu} \simeq 10^8$ which ensures that quantum statistical effects are negligible. ✧

It must be pointed out that our considerations have been based entirely on the independent particle model. Our conclusions turn out to be reliable over a much wider range of conditions than might be expected. After all, the effects of quantum degeneracy appear at high densities and low temperatures—conditions where interactions between particles would seem paramount and where thermal motions are minimal.

In the case of fermions our good fortune may be traced to the Pauli exclusion principle. Regardless of how strongly the fermions interact, the exclusion principle limits the population to one per phase-space cell. Fermions which are free to respond to external forces (those near the Fermi surface) have kinetic

energies far in excess of the thermal energy. Their high kinetic energy enables them to ignore interparticle forces and imitate the motion of nearly free particles.

....

17.6.2 The Fermi energy ε_F of a degenerate gas of free electrons is given by (17.16). As was shown in Example 17.1, the mean kinetic energy of such electrons is $(3/5)\varepsilon_F$. The Coulomb potential energy of two electrons separated by a distance r is e^2/r. Taking $(V/N)^{1/3} = n^{-1/3}$ as a measure of the mean distance between charges suggests that $e^2 n^{1/3}$ is a measure of the mean potential energy of an electron. Provided $(3/5)\varepsilon_F/e^2 n^{1/3} \gg 1$ we expect that the electrons will behave like free particles. This inequality will prevail at sufficiently large n because the potential energy is proportional to $n^{1/3}$ while the kinetic energy is proportional to $n^{2/3}$.

(a) Show that $(3/5)\varepsilon_F/e^2 n^{1/3} = (6\pi^2/5)(3/8\pi)^{2/3} a_0 n^{1/3}$, where $a_0 = \hbar^2/me^2 = 5.29 \times 10^{-9}$ cm is the Bohr radius.

(b) Evaluate the ratio for the typical metallic density $n = 10^{22}/\text{cm}^3$.

....

The existence of two versions of quantum statistics may cause the student to wonder out loud about the range of validity of thermodynamics. We have demonstrated how QSM reduces to CSM at low densities and high temperatures. Is it necessary to revise the laws of thermodynamics to encompass quantum phenomena? The answer is "no"; the old laws will do nicely. Distinctions between QSM and CSM are felt through different expressions used to evaluate thermodynamic properties. For example, the internal energy of a noninteracting gas of bosons or fermions can be expressed by the general relation

$$U = \sum \varepsilon_i N_i$$

where N_i is the number of particles with energy ε_i; N_i depends on T, V, and perhaps other thermodynamic variables. For fermions, N_i is given by (17.13); for bosons N_i is given by (17.21). Despite the fact that the form of $U(T, V, \ldots)$ is different for fermions and bosons, the first law of thermodynamics still reads $dQ = dU + dW$. Likewise, the vast arrays of thermodynamic formulas, such as Maxwell's relations, remain valid regardless of whether one deals with fermions, bosons, or mixtures thereof.

REFERENCES

The story of Planck's struggle with quantum theory is delightfully told in:

M. J. Klein, Thermodynamics and Quanta in Planck's Work, *Phys. Today* **19**, 23 (1966).

Another view of the birth of quantum theory is given in:

K. K. Darrow, The Quantum Theory, *Sci. Am.* **186** (March, 1952). (Also reprinted as *Sci. Am. Offprint* #205.)

The hotel analogy was inspired by a modern classic:

J. D. Trimmer, Hotel Management in Ergodia, *Phys. Today* **15**, 28 (November, 1962).

For alternate derivations of the QSM distributions see:

R. D. Cowan, Derivation of the Energy—Distribution Formulas in Quantum Statistics, *Am. J. Phys.* **25**, 463 (1957)

and

K. K. Darrow, The New Statistical Mechanics, *Bell Sys. Tech. J.* **22**, 362 (1943).

For a careful discussion of the statistical interpretation of the entropy and the revelation which shows h^3 to be the volume of a single phase-space cell consult:

K. K. Darrow, Entropy, *Bell Sys. Tech. J.* **21**, 51 (1942).

Elementary discussions of quantum statistics can be found in several texts. Two which circumvent the quantum mechanical wave function approach and include applications are:

F. W. Sears, "An Introduction to Thermodynamics, The Kinetic Theory of Gases, and Statistical Mechanics," Chapter 16. Addison-Wesley, Reading, Massachusetts (1953).

and

W. Kauzmann, "Thermodynamics and Statistics," Chapter 4. Benjamin, New York (1967).

Chapter 18

Applications of Quantum Statistical Mechanics

This chapter is the quantum mechanical counterpart of Chapter 16. The treatment of the various topics is uneven; at points it is necessary to substitute description for quantum mechanics.

In Section 18.1 we take a detailed look at the theory of blackbody radiation. Recently this oldest of quantum theories has attracted attention in connection with experiments which test the merits of various cosmological hypotheses—notably the "big bang" and "continuous creation" theories. In Section 18.2 we describe the three basic mechanisms by which atoms interact with radiation and then show how these processes are capable of producing both the blackbody spectrum and the spectacular laser action.

Section 18.3 provides an introduction to the physics of neutron stars and white dwarfs—two of the three possible endpoints of stellar evolution. Sections 18.4 and 18.5 revert to the frigid world of dilution refrigerators and superconductivity. The chapter closes with the revelation that there is an upper limit for temperature—an ultimate temperature which cannot be exceeded regardless of how much energy is fed into a physical system.

18.1 BLACKBODY RADIATION, BIG BANG, AND THE PRIMEVAL FIREBALL

Electromagnetic radiation exhibits particle-like properties. A stream of radiation may be treated as a photon gas. The integral spin of a photon brands

it a boson. Within the framework of nonrelativistic quantum theory the photons do not interact with each other. Accordingly, they constitute a gas of independent bosons and their phase-space distribution is given by (17.24). The energy ε of a photon is related to the frequency ν of its source by Planck's famous relation

$$\varepsilon = h\nu$$

The linear momentum of a photon follows from the relation furnished by the special theory of relativity. For a particle of rest mass M, the linear momentum p and energy E are related by (c is the speed of light)

$$E^2 = p^2c^2 + (Mc^2)^2$$

For entities such as the photon, whose rest mass is zero, this reduces to $E = pc$. Thus the linear momentum of a photon p is given by

$$p = \frac{\varepsilon}{c} = \frac{h\nu}{c} \tag{18.1}$$

....

18.1.1 The special theory of relativity shows that the linear momentum of a particle of rest mass M and speed v has the magnitude

$$p = \frac{Mv}{\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}}$$

Use this to prove that any particle with a nonzero linear momentum which has zero rest mass (that is, which does not exist as an object at rest) necessarily travels at the speed of light c.

....

The blackbody spectrum is continuous—all frequencies are present. This situation is handled by writing the Bose–Einstein distribution (17.24) as

$$dN = \frac{dQ}{e^{\beta(\varepsilon - \mu)} - 1} \tag{18.2}$$

where dQ is the number of phase-space cells with energies in the range ε to $\varepsilon + d\varepsilon$ and dN is the corresponding number of photons. Because of the relation $\varepsilon = pc$, we can readily determine the range of momenta corresponding to any range of energies. The number of phase-space cells dQ equals the volume of phase space corresponding to the prescribed range of momenta, divided by the volume of one cell

$$dQ = \frac{\text{(volume of phase space corresponding to momenta in } p \text{ to } p + dp)}{\text{(volume of one phase-space cell)}}$$

The spatial distribution of blackbody photons in a cavity is uniform and isotropic. If V denotes the total volume occupied by the photon gas, we can write for dQ

$$dQ = 2\frac{4\pi p^2\, dp\, V}{h^3} \tag{18.3}$$

We are permitted to use V as the spatial volume element because of the uniformity. The isotropy of the radiation allows us to employ a spherical momentum space volume element $4\pi p^2\, dp$. The volume of one phase-space cell is h^3. The factor 2 displayed so prominently in (18.3) arises from the fact that the spin angular momentum vector of a photon must be parallel or antiparallel to its linear momentum vector. For the prescribed range of p there are $4\pi p^2\, dp/h^3$ cells open to photons having their spin vector parallel to their linear momentum. An equal number of cells can accommodate photons with antiparallel spin and linear momentum vectors. Using (18.1) we can express dQ as

$$dQ = \frac{8\pi V}{c^3}\nu^2\, d\nu \tag{18.4}$$

The chemical potential of a photon gas (ph.g.) is zero,

$$\mu_{\text{ph.g.}} = 0$$

This is related to the fact that the absorption and emission of radiation require an overall conservation of energy, but not photon conservation. For example, an ultraviolet photon could be absorbed and several infrared photons emitted. The chemical potential was introduced as a factor which guaranteed that all candidates for most probable distribution were composed of the same number of particles. Because the total number of photons need not be conserved, the constraint $N = \sum_i^M N_i$ cannot be imposed on the photon gas, with the result that $\mu = 0$. With these results for μ and dQ we can write

$$dN = \frac{8\pi V/c^3}{e^{\beta h\nu} - 1}\nu^2\, d\nu \tag{18.5}$$

The energy of each of the dN photons is $\varepsilon = h\nu$, so that the blackbody energy contained in the frequency range from ν to $\nu + d\nu$ is

$$dU = h\nu\, dN = \frac{8\pi hV/c^3}{e^{\beta h\nu} - 1}\nu^3\, d\nu \tag{18.6}$$

The spectral energy density $u(\nu)$ is defined such that $u(\nu)\, d\nu$ is the energy density (energy per unit volume) in the frequency range ν to $\nu + d\nu$.

Evidently,

$$u(\nu)\,d\nu = \frac{dU}{V} \tag{18.7}$$

giving

$$u(\nu)\,d\nu = \frac{8\pi h}{c^3}\frac{\nu^3\,d\nu}{e^{\beta h\nu}-1} \tag{18.8}$$

The total energy density $u = U/V$ is obtained by integrating $u(\nu)\,d\nu$ over all frequencies.

$$u = \frac{8\pi h}{c^3}\int_0^\infty \frac{\nu^3\,d\nu}{e^{\beta h\nu}-1} \tag{18.9}$$

Setting $x = \beta h\nu$ and recalling that $\beta = 1/kT$, gives

$$u = \left(\frac{8\pi k^4}{h^3c^3}\right)T^4\int_0^\infty \frac{x^3\,dx}{e^x-1}$$

The value of the integral is $\pi^4/15$ (see Example 14.4). The blackbody energy density is proportional to T^4

$$u = aT^4 \tag{18.10}$$

The blackbody energy density constant is

$$a = \frac{8\pi^5k^4}{15h^3c^3} = 7.56\times 10^{-15}\ \ \frac{\text{erg}}{\text{cm}^3\text{-K}^4}$$

....

18.1.2 The energy density of starlight is approximately 1 eV/cm^3. What is the corresponding blackbody temperature?

....

As shown below in Example 18.1, the energy density u and the surface flux are related by

$$F = \frac{c}{4}u \tag{18.11}$$

where c is the speed of light. This result leads to the Stefan–Boltzmann Law (17.1), with $\sigma = ca/4$

$$F = \sigma T^4 \tag{17.1}$$

$$\sigma = \frac{2\pi^5k^4}{15h^3c^2} = 5.67\times 10^{-5}\ \ \frac{\text{erg}}{\text{sec-cm}^2\text{-K}^4}$$

The differential form of (18.11) also holds, the monochromatic emissive power $e(\nu)$ and the spectral energy density being related by

$$e(\nu) = \frac{c}{4}u(\nu) \tag{18.12}$$

Inserting (18.8) for $u(\nu)$ gives

$$e(\nu) = \frac{2\pi h}{c^2}\frac{\nu^3}{e^{\beta h\nu}-1} \tag{18.13}$$

which is Planck's form (17.4), with

$$A = \frac{2\pi h}{c^2}, \qquad B = \frac{h}{k}$$

Figure 18.1 plots $e(\nu)$ versus ν for several different blackbody temperatures.

....

18.1.3 Show that the number density of blackbody photons is given by

$$n = 8\pi\left(\frac{kT}{hc}\right)^3 \int_0^\infty \frac{x^2\,dx}{e^x-1}$$

The value of the integral is 2.404 (see Problem 14.4.4).

....

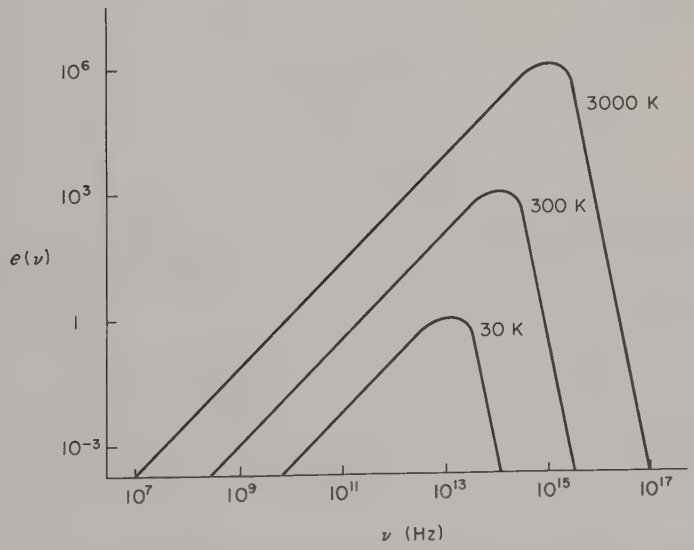


Figure 18.1. The blackbody spectrum for temperatures of 30 K, 300 K, and 3000 K. Both scales are logarithmic, giving the spectrum an appearance different from the version depicted in Figure 17.1. The vertical scale indicates only relative values of $e(\nu)$.

Example 18.1 Blackbody Surface Flux and Energy Density

Theoretical considerations favor the use of the spectral energy density $u(\nu)$. Experimentally, one measures the emissive power $e(\nu)$. The relation between the two is given by (18.12). To derive this, recall the result of Chapter 13, where we showed that the flux of any kinetic quantity $\Phi(\mathbf{v})$ is given by

$$\mathbf{S}_\Phi = \frac{1}{V}\int_{\substack{\text{all}\\ \text{veloc.}}} \Phi(\mathbf{v})\mathbf{v}f(\mathbf{v})\, d^3v$$

To adapt this to the photon gas we replace $f(\mathbf{v})d^3v$ by dN, the number of photons with momenta in a prescribed range. An integration over all phase space replaces the integration over velocities. Since we want the energy flux, we take $\Phi = \varepsilon = h\nu$. Each photon moves at the same speed c and the integration over phase space boils down to integrations over frequency and direction. Let the positive z axis define the outward normal to the radiating blackbody surface. (See Figure 13.3 for the appropriate geometry.) With this choice of coordinates, only the z component of the energy flux will be nonzero, its value being the surface flux F. The z component of the photon velocity is $c\cos\theta$ so that

$$F = \frac{1}{V}\int h\nu c\cos\theta\, dN(\nu,\theta)$$

For dN we write

$$dN = \frac{dQ(\nu,\theta)}{e^{\beta h\nu}-1}$$

with

$$dQ(\nu,\theta) = \frac{2V}{c^2}\nu^2\, d\nu \sin\theta\, d\theta\, d\varphi \tag{18.14}$$

in place of (18.3). The factor $\sin\theta\, d\theta\, d\varphi$ cannot be replaced by 4π here, since photons leave the surface only over the range $\theta = 0$ to $\theta = \pi/2$ (compare Figure 13.3). The integration over phase space gives

$$F = \left\{\frac{2h}{c^3}\int_0^\infty \frac{\nu^3\, d\nu}{e^{\beta h\nu}-1}\right\}\left\{c\int_0^{2\pi} d\varphi \int_0^{\pi/2}\cos\theta\sin\theta\, d\theta\right\}$$

The second factor produces

$$c\int_0^{2\pi} d\varphi \int_0^{\pi/2}\cos\theta\sin\theta\, d\theta = 4\pi\left(\frac{c}{4}\right)$$

so that

$$F = \frac{c}{4}\left\{\frac{8\pi h}{c^3}\int_0^\infty \frac{\nu^3\, d\nu}{e^{\beta h\nu}-1}\right\} \tag{18.15}$$

Comparison of this with (18.9) reveals the desired result

$$F = \frac{c}{4}u \tag{18.11}$$

Further, since

$$F = \int_0^\infty \varepsilon(\nu)\,d\nu, \qquad u = \int_0^\infty u(\nu)\,d\nu$$

it is obvious from (18.15) that the differential relation

$$\varepsilon(\nu) = \frac{c}{4}u(\nu) \tag{18.12}$$

also holds.

It is often desirable to express the energy density in terms of wavelength λ rather than frequency. This is accomplished by introducing a quantity $\Psi(\lambda)$ such that $\Psi(\lambda)\,d\lambda$ is the energy per unit volume contained in the wavelength range from λ to $\lambda + d\lambda$. We write

$$\Psi(\lambda)\,d\lambda = u(\nu)\,d\nu$$

and use the relations†

$$\nu = \frac{c}{\lambda}, \qquad d\nu = \frac{c\,d\lambda}{\lambda^2}$$

to obtain

$$\Psi(\lambda) = \frac{8\pi ch}{\lambda^5}\frac{1}{e^{\beta hc/\lambda} - 1} \tag{18.16}$$

....

18.1.4 (a) Derive Wien's displacement law, which states that the wavelength λ_{max} for which $\Psi(\lambda)$ is a maximum is inversely proportional to the blackbody temperature. Wien advanced the displacement law in 1896, as the empirical relation

$$\lambda_{max} T = 0.294 \quad \text{cm-K}$$

In the course of the solution you will encounter a transcendental equation which may be put in the form

$$(5 - y) = 5e^{-y}$$

---

† We denote by $d\nu$ and $d\lambda$ the magnitudes of the frequency and wavelength ranges; the formal act of differentiating $\nu = c/\lambda$ gives $d\nu = -c\,d\lambda/\lambda^2$, the minus sign indicating that changes in ν and λ are of opposite sign.

Inspection reveals that the solution is nearly $y = 5$. Taking $y = 5$ in the exponential, but not in the factor $(5-y)$, gives a better approximation. Show that

$$\lambda_{\max} T = \frac{ch}{4.97k}$$

and use the tabulated values of c, h, and k to evaluate the right side.

(b) It is a perhaps surprising fact that the frequency $\nu_{\max}$ at which $u(\nu)$ has its maximum value does not correspond to the wavelength $\lambda_{\max}$. That is

$$\nu_{\max} \lambda_{\max} \neq c$$

Determine $\nu_{\max}$ and verify this fact.

18.1.5 A typical room temperature is near 300 K. Would you expect radiation of the corresponding $\lambda_{\max}$ to be visible? Explain, without benefit of numerical calculation. The human eye is most sensitive to wavelengths near 5600 Å. What blackbody temperature corresponds to a $\lambda_{\max}$ of 5600 Å? The Sun emits like a blackbody with a temperature of about 5800 K (over a wide range of wavelengths, including the visible spectrum). What is the corresponding $\lambda_{\max}$? Can you think of any reason for the near coincidence of the two wavelengths?

•••••• ••••••

It was the theory of blackbody radiation which prompted the birth of quantum theory. Recently, experimental measurements of blackbody radiation have helped to settle a question of truly cosmic proportions—the early history of our universe. Until recently two cosmological theories were foremost. The big bang theory was advanced and artfully exposited by Gamow. The steady-state theory was shaped by Bondi, Gold, and Hoyle.

The big bang theory holds that our universe evolved explosively from a superdense state. In its earliest stages the universe was little more than an expanding fireball, inside of which matter was completely dominated by radiant energy—blackbody radiation.

The steady-state theory contends that the large scale observable properties of the universe do not change with time. Observation has shown the universe to be expanding. The mass density is among the observable properties of the universe. For the density to remain constant while the universe expands, there must be a continuous creation of matter. Matter must be created out of nothing to maintain a constant ratio of mass to volume.

Both theories were equally adept at explaining the meager experimental data until recent blackbody radiation measurements ruled in favor of the big bang theory. We can discuss this fascinating topic only briefly. An authoritative

account of the recent experiments is given by Peebles and Wilkinson. A thorough discussion of the experiments and their interpretations is presented by Peebles in his monograph "Physical Cosmology." Brief discussions of the theories, written by Gamow and Hoyle, are included in "The Universe" (Newman, 1957).

By detecting remnants of the primeval fireball, the experiments rule against the steady-state theory. The steady-state cosmology does not contain a plausible mechanism for producing the observed blackbody radiation. The universe envisioned by the steady-state theory was never in a dense state—it must have always been the way we find it today.

The experimental evidence for the primeval fireball, and thus for the big bang theory, is most impressive. The fireball remnant completely fills the universe. In a sense the universe is the granddaddy of all radiation cavities. One of the characteristics of blackbody radiation is its isotropy—the intensity of the radiation is the same in all directions. The observations have confirmed the isotropy of the fireball radiation. There is a slight anisotropy in the observed radiation. This has been attributed to the Doppler effect resulting from the motion of the Earth about the center of our Milky Way galaxy. An even more crucial test is whether or not the spectrum of the radiation fits that of a blackbody. It is here that the experiments appear to clinch the case for the fireball hypothesis. Gamow's 1948 theory of the big bang predicted a blackbody temperature of about 10 K. Subsequent theoretical estimates placed the temperature in the 1–30 K range. Measurements of the fireball radiation have been made at wavelengths ranging from a few millimeters to several centimeters. The measured intensities conform nicely with the emissive power $e(\nu)$ of a 2.7 K blackbody.

Several direct measurements fall in the so-called radio window of our atmosphere. At wavelengths longer than about 10 cm electromagnetic noise from our galaxy obscures the fireball signal. Radiation produced in our own atmosphere and radiation scattered by interstellar dust grains hamper measurements at shorter wavelengths. However, an indirect method has given a fourth point on the fireball spectrum (see Figure 18.2) at a wavelength of 2.6 mm. Because the fireball radiation pervades the universe, the vast interstellar gas clouds (along with everything else in the universe) are swimming about in a 2.7 K ocean of radiation. Among the constituents of these gas clouds is the diatomic cyanogen molecule (CN). These molecules are capable of absorbing radiation of many wavelengths, one of which is 2.6 mm. When a CN molecule absorbs a photon corresponding to the 2.6 mm wavelength its (quantized) energy is increased slightly. In the jargon of quantum theory, the molecule is raised from the ground state to its first excited state. A definite fraction of the CN molecules will be in this excited state. The fraction depends directly on the everpresent Boltzmann factor $e^{-\varepsilon/kT}$ and thus on the temperature. What

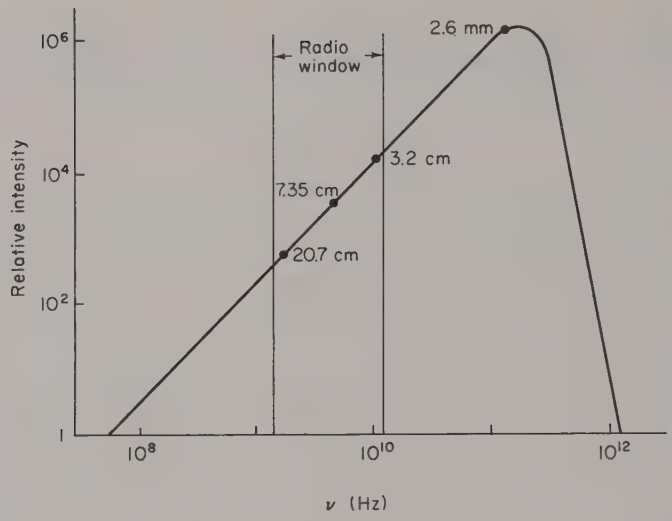


Figure 18.2. The relative intensity for a 2.7 K blackbody plotted versus frequency. Both scales are logarithmic. The relative intensity may be interpreted as the ratio $e(\nu)/e(\nu_0)$, where ν_0 is chosen for convenience. The experimental data points are labeled according to wavelength. The vertical lines define the radio window.

temperature?—the temperature of the CN molecule surroundings, which in this case is the 2.7 K radiation bath. The CN molecules produce certain *absorption* lines in stellar spectra, that is, they intercept certain wavelengths which would otherwise reach the astronomer's photographic plates. The wavelengths absorbed by CN molecules in the ground state are slightly different than those absorbed by molecules in the excited state. By measuring the relative intensities of the absorption lines astronomers can infer the fraction of CN molecules in the excited state. This in turn allows the temperature of the radiation bath to be computed. The experimental results are consistent with a fireball temperature of 2.7 K.

....

18.1.6 The equation of state of a blackbody photon gas is

$$PV = \tfrac{1}{3}U$$

(a) Use this in conjunction with the result of (18.10), $U = aT^4V$, to show that the entropy of such a gas is

$$S = \frac{4}{3}\frac{U}{T}$$

[*Hint:* Start with $dS = dQ/T$, and use the first law of thermodynamics. In an

expanding universe, the density of ordinary matter decreases inversely as the volume ($\rho = m/V$). The corresponding energy density (mc^2/V) also decreases proportional to V^{-1}.]

(b) Show that the blackbody energy density decreases as the inverse 4/3 power of V. [*Hint:* Decide what type of thermodynamic expansion is involved. This result is of importance in understanding the big bang theory. Initially the blackbody energy density far outweighed that of the matter. However, the more rapid decline with volume of the former eventually (about 250 million years after the expansion began) gave ordinary matter the upper hand, and it became possible for galaxy formation to begin.]

18.1.7 The intensity of a typical flashlight beam is roughly 1 W/cm^2. How does this compare with the present fireball energy flux σT^4?

18.1.8 The estimated average mass density in the universe is about 7×10^{-31} gm/cm^3. Cosmological models based on Einstein's general relativity theory require a density of at least 2×10^{-29} gm/cm^3 to give a closed universe. On the basis of present observations the observed expansion of the universe should increase indefinitely. According to Einstein's famous mass–energy relation, $E = mc^2$, any form of energy has a certain mass equivalent. The energy density associated with the primeval fireball is therefore equivalent to a certain mass density. Compute this equivalent mass density. Is it a significant fraction of the value 7×10^{-31} gm/cm^3?

18.1.9 (a) Explain how you could estimate the temperature of the Sun using only your eyes and a thermometer. [You are not permitted to know the Earth–Sun distance! You are permitted to know the angular diameter of the Sun ($\simeq \frac{1}{2}$ degree).]

(b) Make such an estimate. [*Hint:* Assume that the Earth is in radiative equilibrium with the Sun, and that both are blackbodies. Note that the effective absorbing area of the Earth is only $\frac{1}{4}$ of its radiating area.]

....

18.2 STIMULATED EMISSION: THE LASER

Consider a collection of identical atoms capable of absorbing and emitting radiation of frequency ν. The atoms are situated in a blackbody radiation field with which they exchange photons. Each atom has open to it a discrete set of energy states, the quantum energy levels. An atom can undergo a transition—change its energy state—by absorbing or emitting energy in the form of photons. Absorption of a photon increases the energy of an atom. Photon emission

decreases the energy of the atom. Such emissions and absorptions by the atoms represent gains and losses for the blackbody radiation field. At equilibrium, the rates of absorption and emission are equal. It was Niels Bohr who first pictured the emission of radiation as the transition of an atom from a state of higher energy ε_2 to a state of lower energy ε_1. The energy difference $\varepsilon_2 - \varepsilon_1$ is carried away by the emitted photon, it being assumed that the energy and frequency of the photon are related by Planck's relation

$$\varepsilon_{\text{ph}} = h\nu$$

Thus the Bohr theory assumes

$$\varepsilon_2 - \varepsilon_1 = h\nu$$

Einstein recognized three radiative processes: induced absorption, induced emission, and spontaneous emission. Figure 18.3 illustrates the action of the three processes. Einstein discovered the balance existing between these processes by requiring that the energy density of the blackbody radiation field be given by Planck's formula. Figure 18.3a suggests the action of induced absorption. Atoms of energy ε_1 find themselves "bathing" in a stream of

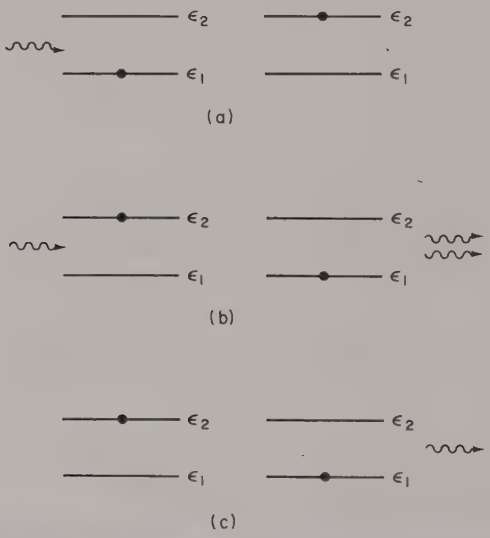


Figure 18.3. The three basic mechanisms by which atoms interact with photons. (a) Absorption: A photon of energy $\varepsilon_2 - \varepsilon_1$ is absorbed by an atom in the level of energy ε_1, raising it to the level of energy ε_2. (b) Induced emission: A photon of energy $\varepsilon_2 - \varepsilon_1$ encounters an atom in the upper level. The incident photon triggers the birth of its "twin." The two photons travel off together, leaving the atom in the lower level. (c) Spontaneous emission: an atom in the upper level spontaneously emits a photon of energy $\varepsilon_2 - \varepsilon_1$. The atom is left in the lower level.

photons. Photons of energy $h\nu = \varepsilon_2 - \varepsilon_1$ can be absorbed, thereby raising the absorbing atoms to the higher energy level. The inverse of induced absorption is induced emission. A stream of photons of energy $h\nu$ may coax atoms of energy ε_2 to emit such photons. As Figure 18.3b suggests each such transition increases the number of photons by one and leaves the atom in the lower energy state.

The rate of induced absorption (emission) is proportional to the number of atoms in the state of energy ε_1 (ε_2) and the density of photons corresponding to frequency ν. This photon density in turn is proportional to the spectral energy density $u(\nu)$. The rates for induced absorption (emission) may therefore be written

$$\frac{\text{number of induced absorptions}}{\text{sec}} = B_{12} N_1 u(\nu)$$

$$\frac{\text{number of induced emissions}}{\text{sec}} = B_{21} N_2 u(\nu)$$

where N_1 and N_2 are the numbers of atoms with energies ε_1 and ε_2 and B_{12} and B_{21} are the Einstein coefficients for induced absorption and emission. The emission of radiation can also proceed spontaneously (Figure 18.3c). Even without the coaxing radiation of frequency ν, an atom of energy ε_2 can "decide" to emit a photon and drop to the state of lower energy. The rate of spontaneous emission is unaffected by the surroundings. It depends only on the internal structure of the atom and the number of atoms in the state of energy ε_2. The rate of spontaneous emissions may therefore be written

$$\frac{\text{number of spontaneous emissions}}{\text{sec}} = A_{21} N_2$$

where A_{21} is called the Einstein coefficient for spontaneous emission. For equilibrium to prevail, the rates of emission and absorption must balance. Adding the rates of spontaneous and induced emission and equating the sum with the rate of absorption yields

$$B_{21} N_2 u(\nu) + A_{21} N_2 = B_{12} N_1 u(\nu)$$

Solving this for $u(\nu)$ gives

$$u(\nu) = \frac{\dfrac{A_{21}}{B_{21}}}{\dfrac{B_{12}}{B_{21}} \dfrac{N_1}{N_2} - 1}$$

With the atoms in thermal equilibrium the ratio of their populations is given

by the Boltzmann factor. Since $\varepsilon_2 - \varepsilon_1 = h\nu$ we find

$$\frac{N_1}{N_2} = \frac{\exp(-\beta\varepsilon_1)}{\exp(-\beta\varepsilon_2)} = \exp(\beta h\nu) \tag{18.17}$$

so that

$$u(\nu) = \frac{\dfrac{A_{21}}{B_{21}}}{\dfrac{B_{12}}{B_{21}} e^{\beta h\nu} - 1} \tag{18.18}$$

If this is to agree with the Planck formula (18.8),

$$u(\nu) = \frac{\dfrac{8\pi h\nu^3}{c^3}}{e^{\beta h\nu} - 1} \tag{18.8'}$$

we must have

$$B_{12} = B_{21} \tag{18.19}$$

$$A_{21} = \frac{8\pi h\nu^3}{c^3} B_{21} \tag{18.20}$$

It is still necessary to compute one of the coefficients via quantum theory. In fact, quantum theory renders it possible to compute all three Einstein coefficients which may be inserted into (18.18) to produce the Planck distribution.

The competition between emissions and absorptions results in the Planck distribution only when the participating atoms are distributed according to the Boltzmann factor (18.17). The Boltzmann formula describes an equilibrium distribution of atoms. When the atoms are not in thermodynamic equilibrium the intensity of the radiation may depart sharply from the Planck form. Such nonequilibrium situations prevail in lasers.† A laser produces an intense, nearly monochromatic, and highly directional beam of radiation. More importantly, laser light is *coherent*. The light waves emitted by individual atoms are locked in step—they are in phase.

To achieve laser action it is necessary to create a nonequilibrium distribution of atoms in which stimulated emission greatly outweighs absorption and spontaneous emission. With reference to Figure 18.3, consider the results which follow if we start with the higher energy level ε_2 populated much more heavily than the lower energy level ε_1. A photon released by spontaneous emission liberates a second photon via stimulated emission. This pair of photons in

† The word LASER is an acronym standing for Light Amplification by Stimulated Emission of Radiation.

turn stimulates further emissions. An avalanche of photons quickly builds up.

The important feature of stimulated emission is that the resulting photons travel off "hand-in-hand." That is, the photon released travels off in the same direction as the photon which triggered its emission. From a wave viewpoint we describe this by saying that the incident wave prompts the atom to radiate a wave of the same frequency, traveling in the same direction. The resultant superposition of the two waves differs from the incident wave in only one essential respect—its amplitude is greater. Each stimulated emission strengthens the wave without altering its frequency or direction.

The growing flux of photons is depleted by absorptions. If the two energy levels are populated in accordance with the Boltzmann formula, absorptions match emissions and no photon avalanche develops. When the population of the upper level is greater than that of the lower level, stimulated emission wins out over absorption and laser action becomes possible.

The Boltzmann population ratio may be upset in a variety of ways, two of which will be discussed shortly. The resulting distribution of atoms is often called an inverted population. Producing an inverted population is referred to as pumping since one destroys the Boltzmann distribution by raising atoms from the lower to the higher energy level.

The two essential components of a laser are the active medium and an optical resonator. The active medium contains the atoms or molecules which emit the laser radiation. In addition the active medium may contain other atoms which assist in producing the required population inversion. The optical resonator is simply a pair of mirrors positioned at either end of the active medium. The mirrors enable the laser light to bounce back and forth through the active medium. Each trip through the active medium strengthens the laser beam. One of the mirrors is coated in such a way that about 1% of the incident light is transmitted. This slight leakage through the mirror constitutes the output of the laser.

The first laser action was achieved by T. H. Maiman in 1960. The active medium in Maiman's laser was a ruby crystal. Ruby is aluminum oxide in which a tiny fraction of the aluminum atoms have been replaced by chromium ions. The captivating color of a ruby stems from the fact that the chromium ions absorb visible light strongly except at the red and blue ends of the spectrum. Pumping was performed optically using a xenon flash tube. Figure 18.4 shows a simplified version of the energy levels and transitions for the ruby laser. The flash pumps chromium ions from the ground state 1 into a band of excited states 3. The upper laser level 2 is populated by nonradiative transitions, in which the chromium ions transfer energy to the crystal lattice instead of emitting a photon. If the xenon flash is sufficiently intense, the ground state population falls below that of level 2, whereupon stimulated emission occurs.

The ruby crystal was a cylinder 4 cm in length and $\frac{1}{2}$ cm in diameter. The ends

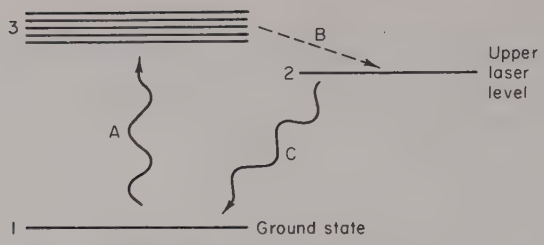


Figure 18.4. Energy levels and transitions for the ruby laser. A, Pump transition $1 \rightarrow 3$; B, nonradiative transition $3 \rightarrow 2$; C, laser transition $2 \rightarrow 1$. (See text for details.)

of the ruby were cut parallel, polished optically flat, and partially silvered, thus forming an optical resonator. Maiman's laser produced an intense pulse of coherent red light at a wavelength of 6943 Å. The pulse lasted about $\frac{1}{2}$ msec. Before moving on to describe the helium–neon laser we must understand those features of the ruby energy levels which promote laser action.

An important consideration in selecting an active medium is the scale of spontaneous emission *lifetimes* of the energy levels involved. Spontaneous emission is the optical counterpart of radioactive decay. While it is not possible to predict when a given radioactive nucleus will decay nor when a given excited atomic state will emit spontaneously, it is possible to establish average lifetimes against such events. For excited atoms typical lifetimes are on the order of 10^{-8} sec. Unusually long lifetimes of a few milliseconds are not uncommon. Such relatively long-lived states are referred to as *metastable*. In ruby the upper laser level 2 is metastable with a lifetime of about 3 msec. The long lifetime against spontaneous emission lets the population of the level build up, helping establish a population inversion. Also helpful is the fact that the nonradiative transitions $3 \rightarrow 2$ occur more rapidly than the radiative transitions which return the chromium ion to its ground state $3 \rightarrow 1$. Thus the $1 \rightarrow 3$ pumping transition populates level 3, which empties into level 2 more rapidly than the inverse $3 \rightarrow 1$ transitions can return the ions to the ground state. The fraction of the ions pumped out of the ground state increases with the intensity of the flash. A sufficiently large flash intensity produces a population inversion and thereby triggers laser action.

A prime drawback of the three-level operation typified by the ruby laser is that the laser transition terminates on the ground state level. A significant fraction of the ions must be pumped from the ground state to achieve an inverted population. A more efficient scheme is offered by four-level operation (see Figure 18.5). The fourth level replaces the ground state as the terminus of the laser transition. Prior to pumping, both laser levels are essentially empty by comparison to the ground state, because of the exponential Boltzmann factor. A weak but selective pumping operation suffices to establish an inverted

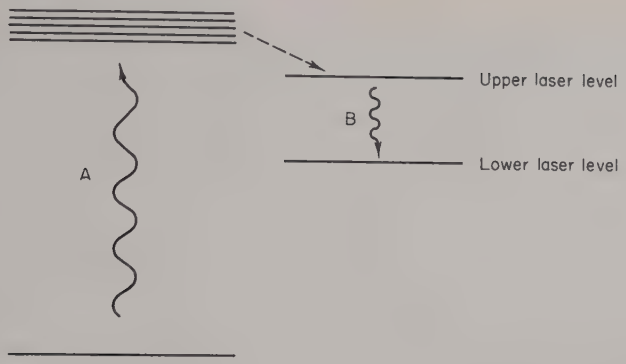


Figure 18.5. A four-level scheme for achieving an inverted population. A, Pump transition; B, laser transition.

population. The pumping action must be effective in populating the upper laser level, but not the lower level. In practice this selectivity is achieved in a variety of ways.

The high powers needed to achieve an inverted population in the three-level ruby laser restricted it to pulsed operation. The four-level scheme makes it feasible to operate a continuous wave (CW) laser. The helium–neon laser is a versatile CW device. It gives the appearance of being a neon sign placed between two mirrors. A popular version of the helium–neon laser operates in the visible spectrum at a wavelength of 6328 Å. A mixture of gaseous helium and neon constitute the active medium. The laser transitions between two levels of the neon atom produce the red 6328 Å radiation. The helium plays a role in the pumping scheme. Figure 18.6 shows the pertinent helium and neon

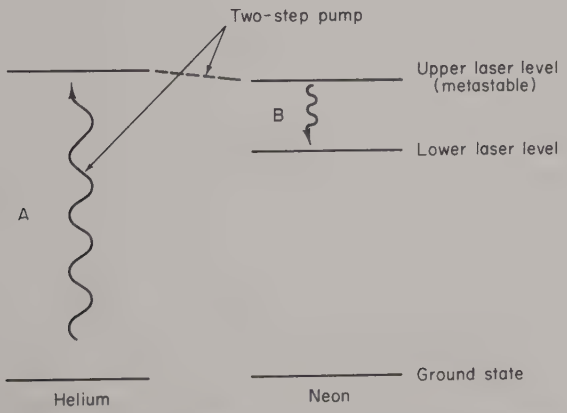


Figure 18.6. Energy levels involved in producing an inverted population in the He–Ne laser. A, Electrical discharge excitation; B, laser transition.

energy levels. An electrical discharge provides energetic electrons which pepper the helium atoms, boosting them into excited states. The energy of the excited state shown for the helium atom is a few hundredths of an electron volt greater than the energy needed to excite the upper laser level of the neon atom. A collision between an excited helium atom and a ground state neon atom can result in a transfer of energy to the neon atom. The neon atom is raised to the upper laser level and the helium atom returns to its ground state. The energy transfer is favored by the very small difference in the excitation energies. The small excess of energy is parceled out as translational kinetic energy between the colliding atoms. This collisional process is selective. The excitation energy of the helium atoms is nearly two electron volts higher than the excitation energy of the lower level of the laser transition. This large difference makes it very unlikely that an excited helium atom will transfer its energy to a neon atom and thereby populate the terminal state of the laser transition.

Of course, collisions between the swift electrons and neon atoms in the ground state will populate both laser levels. However, such collisions are not selective. They do not tend to produce a population inversion. The key to the pumping scheme is the tiny energy difference between the two excited states of helium and neon which favors the upper laser level.

The laser output is highly directional. This results from the fact that the amplitude of the wave builds up via repeated trips through the active medium. Only waves starting out more or less parallel to the axis of the optical resonator will be trapped and have the opportunity to grow in amplitude. Waves starting in other directions are refracted or reflected out of the active medium. The slight divergence of the laser beam is caused by diffraction as the escaping light passes through the mirror. The high directionality of laser light was demonstrated spectacularly in an experiment which beamed laser light to the Moon. The light was reflected back to Earth by an array of corner cubes left by the Apollo 11 astronauts. Optical components expanded the laser beam to give it a diameter of 10 ft when it left Earth. After traveling the nearly 250,000 miles to the Moon its diameter was approximately 1 mile. In traveling to the Moon the beam diameter grew by only 1 ft every 50 miles.

Many different types of lasers have been developed. Technical advances are so swift that any listing of laser applications would quickly become obsolete. Early solid-state lasers, such as the ruby laser, were much more powerful than the gas lasers simply because of the greater density of the active medium in a solid. Gas lasers offered the advantage of continuous operation in contrast to the pulsed action of the solid-state lasers. The hectic pace at which laser research progresses is rapidly eliminating these early drawbacks. There are now CW solid-state lasers. The recently developed carbon dioxide gas laser develops several kilowatts of power—enough to cut through a $\frac{1}{4}$-in. steel plate in a few seconds. Whereas the early lasers operated at a fixed wavelength, recent innova-

tions have made it possible to tune a laser—to select any wavelength within a definite range. Chemical lasers offer another promising area for development. In a conventional laser the active medium draws its energy from an external source. This source may be optical, as in the case of the xenon flash tube, or it may be electrical, as it is with the helium–neon laser. In a chemical laser the active medium draws its energy from a chemical reaction. In fact, the chemical reactions take place inside the active medium. The chemical laser is self-pumping. It is somewhat like having a flashlight in which the bulb and batteries are a one-piece unit.

18.3 WHITE DWARFS, NEUTRON STARS, AND PULSARS

Stars are born when gravitational contraction causes the density and temperature of a protostar to reach levels where thermonuclear reactions are ignited. Initially, the nuclear fuel is hydrogen. Hydrogen burning produces helium and releases energy. Part of this energy remains in the star as the kinetic energy of thermal motion while the rest escapes as radiation, giving the star its luminosity. The pressures associated with the thermal motions and the flow of radiation counterbalance the compressive gravitational force, thereby enabling the star to maintain a stable configuration.

When a nuclear fuel is depleted, gravitational contraction begins anew. Part of the gravitational energy release is used to maintain the stellar luminosity. The remainder serves to raise the temperature, perhaps to a point where the ashes of earlier reactions ignite. The intricacies of stellar burning need not concern us here. After a lifetime on the order of billions of years the stellar furnace runs out of nuclear fuel.

Once a star has exhausted its accessible nuclear fuel the inevitable gravitational contraction resumes. According to currently accepted views this contraction terminates in one of three possible configurations—a white dwarf, a neutron star, or a black hole. The mass of the star as it begins to collapse is the most important quantity in determining which of these stellar corpses results. This need not be the final mass of the star since gravitational collapse may trigger a supernova which blasts a portion of the stellar mass outward into space.

The Sun is likely to die as a white dwarf. Its final gravitational collapse will cause the density to increase to the point where the electrons become degenerate. As degeneracy mounts the electron pressure increases sharply. This rise in pressure slows the gravitational collapse and eventually causes the contraction to subside. A stable configuration results in which the gravitational compression is balanced by the pressure of the degenerate electron gas. There is an upper limit to the mass which can be supported by the electron degeneracy

pressure. This limiting value is about 1.2 times the mass of the Sun† and is known as the *Chandrasekhar limit.*

If the mass of the collapsing star is sufficiently great, gravitation can overwhelm the electron degeneracy pressure and a neutron star may result. The density becomes so great in the interior of a neutron star that atomic nuclei dissolve. A mixture forms in which about 99% of the particles are neutrons, hence the name of the star. The remaining 1% is made up of protons and electrons along with a smattering of more exotic particles. The neutrons are strongly degenerate and it is their pressure which resists the crush of gravity and sustains a stable configuration.

Radio astronomers recently discovered the so-called pulsars. Pulsars are characterized by the periodic emission of intense bursts of electromagnetic energy. The presently accepted thinking is that pulsars are rotating neutron stars which generated intense magnetic fields during gravitational collapse. The pulsar magnetic fields are a million times stronger than any yet produced in a laboratory. This field renders the pulsar atmosphere anisotropic, permitting charged particles and electromagnetic radiation to escape only along certain special directions. Two models, one as plausible and picturesque as the other, have attracted the most attention. One model has the rotation of the star and the channeling action of its magnetic field producing a narrow funnel-shaped flow of radiation (see Figure 18.7). The rotating beams of radiation sweep through space like an airport beacon and periodically illuminate the radio telescopes.

The second model proposes that the radiation is emitted by electrons which leave the stellar surface and are accelerated as they slide along the rotating magnetic field lines. Neither the particle nor the field can rotate at a speed in excess of the speed of light. As the particles approach the speed of light their acceleration results in an intense and highly directional beam of radiation. The beam originates well away from the stellar surface, near the velocity-of-light circle. The velocity-of-light circle is centered at the star and has a radius given by

$$R = \frac{2\pi c}{P}$$

where P is the rotational period of the star and c is the speed of light. Most pulsars have rotation periods on the order of 1 sec. We will see later that this was an essential clue to their identification as rotating neutron stars.

The third possible endpoint of stellar evolution is a black hole. A black hole presumably results when the collapsing mass is so great that not even the

---

† The mass of the Sun (2×10^{33} gm) is a convenient stellar mass unit and is designated as $1M_{\odot}$. The Chandrasekhar limit is approximately $1.2M_{\odot}$.

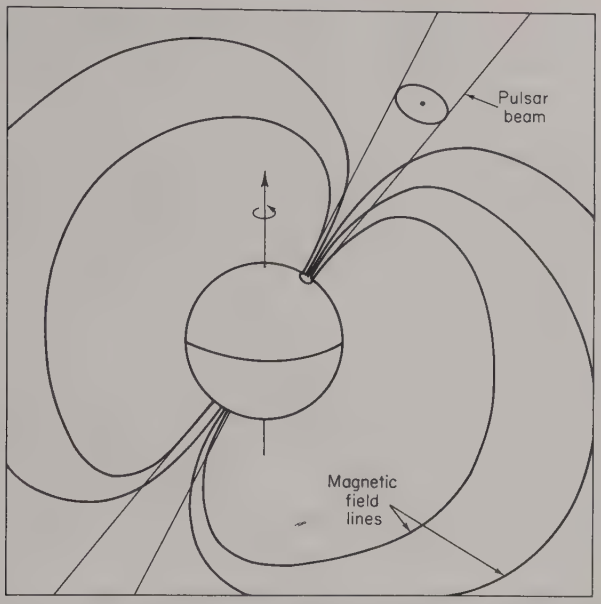


Figure 18.7. The pulsar magnetic field channels radiation, allowing intense beams to escape from the magnetic polar regions. The axis of rotation does not coincide with the magnetic axis. As a result the beams sweep through space like gigantic searchlights.

neutron degeneracy pressure can check its course. The black hole gains its name from the fact that not even light can escape its gravitational pull.

The black hole is presently a speculation, based on the general theory of relativity. None has been detected and it is not yet clear how astronomers might best search for them. Neutron stars have only recently been observed as pulsars. Of the endpoint trio, white dwarfs have the oldest pedigrees. Their discovery early in this century was a major surprise to astrophysicists and an understanding of their structure had to await the invention of quantum statistical mechanics.

Our discussion is limited to white dwarfs and neutron stars.† These objects are alike in that their stability against gravitational collapse is sustained by the pressure of degenerate fermions. This pressure is a direct manifestation of the Pauli exclusion principle.

In the development which follows we obtain estimates for the mass, radius, and magnetic field strength of degenerate stars—white dwarfs and neutron stars. The analysis is the sort physicists are most fond of; it amounts to a

† References at the end of this chapter include articles that discuss gravitational collapse and black holes.

sophisticated dimensional analysis. The technique is not geared for numerical precision. It can tell us, for example, that the mass of a white dwarf is approximately $1M_\odot$ but it can pin down the precise value only to within a factor of 5 or so.

Because of the inherent crudeness of the method we do not hesitate to take minor liberties with numerical quantities. For example, the total number of nucleons (protons or neutrons) in the Sun is 1.2×10^{57}. It is completely within the spirit of our procedures to round this number to 10^{57}. The solar nucleon number is designated by $N_\odot$

$$N_\odot = 10^{57}$$

Another dimensionless number which enters our considerations is $\hbar c/Gm_n^2$, where G is the gravitational constant and m_n is the nucleon mass. The numerical value of this quantity is 1.6×10^{38}, which we round to 10^{38}. A windfall of this minor fudging is the discovery that

$$N_\odot^{2/3} \simeq \frac{\hbar c}{Gm_n^2} \simeq 10^{38}$$

We make frequent use of this near equality in our analysis.

The gravitational potential energy of a star decreases as it collapses. The potential energy sacrificed shows up as an increase in kinetic energy and in other forms of energy, notably radiation. Energy is conserved; gravitational energy is converted into kinetic energy and radiation.

The kinetic energy increase is easily understood. Objects falling under the influence of gravity speed up. However, only about $\frac{1}{2}$ of the gravitational energy is converted into kinetic energy. Much of the remainder escapes from the star as radiation.

If a star of mass M collapses to some final radius R (small by comparison to its initial radius), the gravitational energy released is of order GM^2/R. This figure disregards a numerical factor of order unity whose value depends on how the mass is distributed throughout the star. In order of magnitude then, the kinetic energy increase of the particles is also GM^2/R. Further, the kinetic energy gained during collapse is substantially greater than the initial kinetic energy. We may therefore set

$$\text{kinetic energy of particles in final configuration} = \frac{GM^2}{R}$$

Let N denote the number of degenerate particles whose pressure supports the star against gravitational compression. As degenerate matter these particles are packed as tightly as the exclusion principle permits. The average kinetic energy of such particles is close to $\frac{3}{5}\varepsilon_F$, where ε_F is the Fermi energy (compare Example 17.1). We ignore the factor $\frac{3}{5}$ and adopt $N\varepsilon_F$ as the kinetic energy of

the degenerate matter. Equating $N\varepsilon_F$ with the expression for the kinetic energy deduced above gives

$$\frac{GM^2}{R} = N\varepsilon_F \tag{18.21}$$

To make use of this relation we must express the Fermi energy in terms of R and N. Let p_F denote the Fermi momentum. The representative volume of momentum space per particle is $\frac{4}{3}\pi {p_F}^3$; the average spatial volume per particle is $\frac{4}{3}\pi R^3/N$. The product of these two gives the phase-space volume per particle. Changing the numerical factor slightly for later convenience we can write

$$(2\pi p_F)^3 \frac{R^3}{N} = \text{(phase-space volume per fermion)}$$

The degenerate fermions fill phase space as densely as the exclusion principle permits, namely one per phase-space cell. This means that the phase-space volume per particle equals h^3, the volume of a single phase-space cell,

$$(2\pi p_F)^3 \frac{R^3}{N} = h^3 \tag{18.22}$$

The Fermi momentum follows as

$$p_F = \hbar\left(\frac{N}{R^3}\right)^{1/3} \tag{18.23}$$

Observe that the Fermi momentum is independent of the mass of the fermions; it is completely determined by their number density.

By eliminating p_F and ε_F between (18.21) and (18.23) we obtain a relation between R and N. Since N is directly proportional to the stellar mass this result also establishes the mass–radius relation for the star.

The relation between ε_F and p_F for nonrelativistic fermions is $\varepsilon_F = {p_F}^2/2m$. However, in white dwarfs the electrons move at relativistic speeds—they have kinetic energies comparable to their rest mass energy mc^2. The special theory of relativity shows that the kinetic energy of a particle is given by

$$\sqrt{(pc)^2 + {e_0}^2} - e_0$$

where p is the linear momentum and $e_0 = mc^2$ is the rest mass energy.

....

18.3.1 (a) Use the binomial expansion to show that $\sqrt{(pc)^2 + {e_0}^2} - e_0 \simeq p^2/2m$, in the nonrelativistic limit where $pc \ll e_0 = mc^2$.

(b) What is the ultrarelativistic limit for the kinetic energy ($pc \gg e_0$)?

....

Since we have identified ε_F as the kinetic energy we write

$$\varepsilon_F = \sqrt{(p_F c)^2 + e_0^2} - e_0 \tag{18.24}$$

Inserting this into (18.21) gives

$$\frac{GM^2}{R} = N[\sqrt{(p_F c)^2 + e_0^2} - e_0] \tag{18.25}$$

Minor algebraic manipulation converts this into

$$\left(\frac{GM^2}{RN}\right)^2 + 2e_0\left(\frac{GM^2}{RN}\right) = (p_F c)^2 \tag{18.26}$$

Finally, we use the result of (18.23) to eliminate p_F, obtaining

$$\left(\frac{GM^2}{RN}\right)^2 + 2e_0\left(\frac{GM^2}{RN}\right) = (\hbar c)^2\left(\frac{N}{R^3}\right)^{2/3} \tag{18.27}$$

Once we specify the relation between M and N, (18.27) gives us the sought-for relation between M and R.

For a neutron star, where the neutrons supply the degeneracy pressure and the mass we have

$$M = Nm_n \tag{18.28}$$

where $m_n = 1.67 \times 10^{-24}$ gm is the nucleon mass. In a white dwarf, where electrons supply the degeneracy pressure, (18.28) does not apply because the number of nucleons is not equal to the number of electrons. Let μ_e denote the number of nucleons per electron. In place of (18.28) we have

$$M = \mu_e N m_n \tag{18.29}$$

In a white dwarf it is expected that $\mu_e \simeq 2$. This is the value of μ_e appropriate for a star which has burned its hydrogen into helium.†

In keeping with the spirit of our method we ignore the μ_e factor and use (18.28) throughout. Problem 18.3.3 provides an opportunity to demonstrate that our method yields the correct μ_e dependence.

....

18.3.2 Explain why $\mu_e = 2$ for a pure helium star.

....

† The temperature of a stellar interior is so high that all matter is completely ionized. Thus helium is not present as a neutral atom, but as an α particle and 2 electrons. Likewise, the hydrogen exists as free protons and electrons.

If we insert (18.28) into (18.27), multiply through by $R^2/(Gm_n^2)^2$, and recall that $N_\odot^{2/3} \simeq \hbar c/Gm_n^2$, we obtain

$$N^2 + 2N^{4/3}N_\odot^{2/3}\left(\frac{R/N^{1/3}}{\lambda_c}\right) = N^{2/3}N_\odot^{4/3} \tag{18.30}$$

The quantity

$$\lambda_c = \frac{\hbar}{mc} \tag{18.31}$$

is known as the *Compton wavelength* wherein m is the mass of the degenerate fermion which supplies the pressure. For white dwarfs m is the electron mass; for neutron stars m is the nucleon mass. Numerically,

$$\begin{aligned} \lambda_c &= 3 \times 10^{-11} \text{ cm}, && \text{white dwarf} \\ \lambda_c &= 2 \times 10^{-14} \text{ cm}, && \text{neutron star} \end{aligned} \tag{18.32}$$

Solving (18.30) for $(R/N^{1/3})$ gives the mass–radius relation

$$\left(\frac{R}{N^{1/3}}\right) = \frac{1}{2}\lambda_c\left(\frac{1}{x} - x\right) \tag{18.33}$$

where

$$x = \left(\frac{N}{N_\odot}\right)^{2/3} = \left(\frac{M}{M_\odot}\right)^{2/3} \tag{18.34}$$

The last equality follows from (18.28) and the fact that $M_\odot = N_\odot m_n$ is the mass of the Sun. The quantity $((1/x) - x)$ is negative for values of x greater than one. Since the radius cannot be negative we conclude that x must not exceed unity

$$x = \left(\frac{M}{M_\odot}\right)^{2/3} \leqslant 1 \tag{18.35}$$

Equation (18.35) shows that there is an upper limit to the mass of a degenerate star, namely,

$$M \leqslant M_\odot \tag{18.36}$$

The upper limit of $M_\odot$ corresponds to $x = 1$ and $R = 0$. The idea of a zero radius star is of course not realistic. Nevertheless, the existence of an upper limit has an important implication. It means simply that normal stars of mass substantially greater than $M_\odot$ cannot die as degenerate stars unless they first manage to shed mass. Observations reveal that a sizable fraction of the stars have masses above $M_\odot$ and this poses a question as to their ultimate fate.

The original analysis by B. S. Chandrasekhar gave an upper limit of $5.76M_\odot/\mu_e^2$ for white dwarfs. More refined calculations lowered the Chandrasekhar limit to approximately $5M_\odot/\mu_e^2$. For the expected value of $\mu_e \simeq 2$ the Chandrasekhar limit is about $1.2M_\odot$. Refined calculations, which account for the interaction between nucleons, show that neutron star configurations can exist over the range from about $0.25M_\odot$ to near $2M_\odot$. Thus the scale of one solar mass suggested by our rudimentary analysis is not far off the mark.

....

18.3.3 Show that the limiting mass for a white dwarf emerges as $M_\odot/\mu_e^2$ when (18.29) is used in place of (18.28).

....

For modest values of x, (18.33) shows that the scale for the stellar radius is

$$R \simeq \lambda_c N^{1/3} \tag{18.37}$$

For a star with $M \simeq M_\odot$ we have $N \simeq N_\odot$. Inserting $N_\odot^{1/3} \simeq 10^{19}$ for $N^{1/3}$ and noting (18.32) we find†

$$\begin{aligned} R_{\text{W.D.}} &\simeq 3 \times 10^8 \text{ cm} \simeq 5 \times 10^{-3} R_\odot \\ R_{\text{N.S.}} &\simeq 2 \text{ km} \end{aligned} \tag{18.38}$$

The white dwarf (W.D.) figure is comparable to the radius of the Earth (6×10^8 cm). Precise calculations for white dwarfs indicate somewhat larger radii, the exact value depending on the stellar mass. Calculations for neutron stars (N.S.) predict a radius of approximately 13 km, independent of the stellar mass.

As yet no experimental determinations of mass or radius have been obtained for neutron stars. Over 200 white dwarfs have been identified, but mass and radius determinations have been made for only eight. Data for three white dwarfs are

$$\begin{aligned} &\text{40 Eri B } (R = 1.6 \times 10^{-2} R_\odot; \quad M = 0.44 M_\odot) \\ &\text{Sirius B } (R = 2 \times 10^{-2} R_\odot; \quad M = 0.98 M_\odot) \\ &\text{Procyon B } (R = 10^{-2} R_\odot; \quad M = 0.65 M_\odot) \end{aligned}$$

The data for all eight stars are consistent with our results. Although meager, they reassure us that our primitive approach defines reliable mass and radius scales.

---

† $R_\odot$ is the radius of the Sun 7×10^{10} cm.

.....

18.3.4 Estimate the average mass density of a white dwarf and a neutron star. Compare the latter with the density of an atomic nucleus ($\rho_{nuc} \simeq 10^{15}$ gm/cm^3). At your computed densities, what would be the weights (on Earth) of 1 cm^3 of such matter?

.....

The magnetic field at the surface of a neutron star is expected to be on the order of 10^{12} G. By comparison, the strongest fields produced by superconducting magnets are less than 10^6 G. Such intense stellar magnetic fields presumably arise from flux trapping. The ionized stellar material is a nearly perfect electrical conductor. As a consequence of this the magnetic lines of force are swept inward with the collapse. The magnetic flux threading the star is thereby trapped—it remains constant during the collapse. The magnetic flux Φ is proportional to the product of the magnetic field H and the cross-sectional area of the star πR^2,

$$\Phi \propto \pi R^2 H$$

It follows from the constancy of Φ that

$$R_i^2 H_i = R_f^2 H_f \tag{18.39}$$

where i and f refer to the initial and final configurations. If a star the size of the Sun ($R_i \simeq 10^{11}$ cm), with a field of order 100 G collapses to a neutron star configuration ($R_f \simeq 10^6$ cm), the resulting magnetic field strength is 10^{12} G. Field strengths of this size are consistent with the observed rates of pulsar slow-down, the barely perceptible increases of pulsar periods.

A similar multiplication of the magnetic field strength should accompany the formation of a white dwarf. Recent observations of white dwarf spectra have indicated field strengths in the 10^4–10^5 G range.

.....

18.3.5 The magnetic field at the surface of the Sun is roughly 1 G. The strength of the field in the solar interior is not known with any degree of certainty. However, the field is stronger in the interior than at the surface. Estimate the field strength which would result if the Sun collapsed to a white dwarf of radius $0.01 R_\odot$.

.....

It was noted earlier that most pulsars have a period near 1 sec. The range of pulsar periods was an important clue in their identification as rotating neutron stars. After the discovery of pulsars in 1967 both neutron stars and white dwarfs were suggested as possible sources of the radiation. Both rotation and vibration were advanced as the mechanisms responsible for the clock-like

regularity of pulsar signals. White dwarfs and vibrating neutron stars were shown to lack one or more of the necessary characteristics, leaving the rotating neutron star as the only likely candidate.

The rotating white dwarf was eliminated on dynamic considerations. The gravitational acceleration at the surface of a star of radius R and mass M is $g = GM/R^2$. The minimum period of rotation ($P_{\min}$) of such a star is fixed by Newton's second law

$$g = R\omega^2, \qquad \omega = \frac{2\pi}{P_{\min}}$$

Periods less than $P_{\min}$ would result in the star tearing itself to shreds. Gravity would no longer be able to supply the necessary centripetal acceleration. From $g = R\omega^2$ we find

$$P_{\min} = 2\pi \frac{R^{3/2}}{\sqrt{GM}} \tag{18.40}$$

For a neutron star one finds

$$P_{\min} \simeq 10^{-3} \text{ sec}$$

while for a white dwarf

$$P_{\min} \simeq 10 \text{ sec}$$

The observed pulsar periods range from 33 msec to about 3 sec, thereby eliminating the rotating white dwarf. A white dwarf is simply too big to rotate once each second.

Vibrating white dwarfs and neutron stars also may be ruled out as likely seats of pulsar activity. The fundamental period of pulsation of a star is proportional to the quotient of the stellar radius and the characteristic soundspeed (V_s)

$$P \sim \frac{R}{V_s} \tag{18.41}$$

The vibrations propagate as pressure variations. Hence the soundspeed in (18.41) is the speed characteristic of the degenerate particles which supply the stellar pressure.

In a neutron star the pressure is so enormous that the soundspeed approaches a fundamental limit—the speed of light. With $R \simeq 10^6$ cm and $V_s \simeq 10^{10}$ cm/sec the fundamental pulsation period of a neutron star is

$$P_{\text{N.S.}} \simeq 10^{-4} \text{ sec}$$

The somewhat less rigid structure of white dwarfs propagates sound at speeds

of order 10^8 cm/sec. With $R \simeq 10^8$ cm, the fundamental pulsation period for white dwarfs is

$$P_{\text{W.D.}} \simeq 1 \text{ sec}$$

The pulsar located in the Crab nebula has a period of 1/30 sec. The Vela pulsar period is less than 1/10 sec. These pulsars cannot be vibrating white dwarfs. Conversely, pulsating neutron stars are ruled out because their fundamental period is much too low.

The remaining pulsar candidate—the rotating neutron star—is left as the only serious contender. Indeed, pulsar models based on the neutron star configuration have successfully explained many of the observations.

Problems associated with the structure of neutron stars provide a fertile field for applications of statistical physics. We conclude our discussion of degenerate stars by showing how one arrives at the conclusion that a neutron star is a blend of 99% neutrons and 1% electrons and protons.

The free neutron is an unstable particle and decays into a proton, electron, and antineutrino,

$$\text{n} \rightarrow \text{p} + \text{e}^- + \tilde{\nu} \tag{18.42}$$

The rest mass energy of a neutron exceeds that of the products. The energy difference shows up as the kinetic energy of the decay products. If an electron and proton collide with sufficient energy, the inverse reaction can occur

$$\text{p} + \text{e}^- \rightarrow \text{n} + \nu \tag{18.43}$$

In a neutron star interior the Fermi energy of the electrons is sufficiently great that this inverse reaction is energetically possible. An equilibrium is established, the rate of neutron decay being balanced by the rate of neutron production.

The condition for chemical equilibrium was established in Chapter 8 (see Example 8.3). The balance between neutron formation and decay is subject to the same conditions. In particular, (8.47) applies to (18.42). Equation (8.47) describes the balance of chemical potentials. For (18.42) it becomes

$$\mu_{\text{n}} = \mu_{\text{p}} + \mu_{\text{e}} + \mu_{\tilde{\nu}} \tag{18.44}$$

The chemical potential for the antineutrinos is zero. For the other species there is only an insignificant difference between the chemical potential and the Fermi energy. Thus (18.44) may be replaced by

$$\varepsilon_{\text{n}} = \varepsilon_{\text{p}} + \varepsilon_{\text{e}} \tag{18.45}$$

Electrical neutrality demands equal numbers of electrons and protons. As was demonstrated earlier, the Fermi momentum depends on the number density but not on the mass of the fermion. The electrons and protons therefore have equal Fermi momenta. By virtue of their smaller masses the electrons acquire

a much higher Fermi energy than do the protons. Ignoring ε_p in favor of ε_e in (18.45) gives

$$\varepsilon_n \simeq \varepsilon_e \tag{18.45a}$$

The neutrons move with nonrelativistic energies, whereas the electrons are ultrarelativistic (see Problem 18.3.6). We have

$$\frac{p_n^2}{2m_n} = \varepsilon_n, \qquad p_e c = \varepsilon_e$$

where p_n and p_e are the Fermi momenta. The Fermi momentum follows from the "pack-'em-in-as-tightly-as-Pauli-allows" principle,

$$p_F = \hbar(3\pi^2 n)^{1/3} \tag{17.17'}$$

where n is the number density. The equilibrium condition (18.45a) then becomes

$$\frac{\hbar^2}{2m_n}(3\pi^2 n_n)^{2/3} = \hbar c(3\pi^2 n_e)^{1/3}$$

which converts to

$$\frac{n_e}{n_n} = \frac{3\pi^2}{8}\left(\frac{\hbar}{m_n c}\right)^3 n_n \simeq 10^{-41} n_n \tag{18.46}$$

The mass density in a neutron star is on the order of 10^{15} gm/cm^3 which means the neutron number density is about 10^{39} cm^{-3}. Thus $n_e/n_n \simeq 10^{-2}$; neutrons are roughly 100 times more plentiful than electrons and protons.

....

18.3.6 (a) Take $n_e = 10^{37}$ cm^{-3} and compute the electron Fermi energy (assume they are ultrarelativistic). How does your result compare with the electron rest mass energy $mc^2 = 0.51$ MeV? Was it justifiable to treat the electrons as ultrarelativistic?

(b) Determine the neutron number density and mass density corresponding to $n_e = 10^{37}$ cm^{-3}.

....

18.4 ^{3}He–^{4}He DILUTION REFRIGERATORS

During the past decade a new type of ultra low temperature refrigerator has been developed. It utilizes a mixture of helium isotopes, ^{3}He and ^{4}He. It is referred to as a *dilution refrigerator* because cooling is achieved by diluting the mixture. There are, in fact, two ways of cooling via dilution. Interestingly

enough, the two dilution schemes are analogous to two classic methods of cooling—namely, adiabatic expansion and forced evaporation. To help expose the analogies we first review these two classic coolers. Following this review we catalog the relevant properties of ^{3}He and ^{4}He, and mixtures of the two. The two methods of cooling via dilution are then discussed in some detail. Commercially available dilution refrigerators can maintain continuous temperatures below 0.03 K.

The thermodynamic basis of the adiabatic expansion method is this: A gas expands adiabatically and performs work against its surroundings. The gas imports no heat, hence the work is done at the expense of its internal energy. An internal energy decrease is accompanied by a temperature drop. At the microscopic level one pictures molecules striking the expanding walls of their container. They transfer energy to the moving walls. The decrease in their kinetic energy registers macroscopically as a temperature drop.

Cooling by forced evaporation works like this: Suppose one starts with a liquid in equilibrium with its vapor, and then begins to pump away the vapor. This upsets the thermodynamic equilibrium. Liquid evaporates in an effort to restore the equilibrium, but evaporation requires heat—the latent heat of vaporization. The vapor carries off this latent heat, which is furnished by the liquid. So, in trying to maintain thermodynamic equilibrium with its vapor, the liquid rejects heat and is cooled. Microscopically, we know that temperature measures the average kinetic energy of the molecules. To overcome the attractive forces of its neighbors in the liquid and escape to the vapor, a molecule must have a kinetic energy considerably above the average. To put it briefly, only the fast molecules can escape. Their exodus lowers the average kinetic energy—and thus the temperature—of those which remain in the liquid.

There are inherent limitations on the classic coolers which prevent them from reaching temperatures below about 0.3 K. Adiabatic expansion machines suffer from the obvious difficulty of lubricating moving parts at liquid helium temperatures. Furthermore, the fractional change in temperature $\Delta T/T$ resulting from an adiabatic expansion decreases with the temperature. These drawbacks limit adiabatic expansion coolers to temperatures above about 5 K.

The lowest temperature which can be achieved by forced evaporation is limited by the rate at which the vapor can be pumped away. The pumping speed is proportional to the pressure head, the pressure difference between the vapor and the pump. For practical purposes the pressure head equals the vapor pressure, but the vapor pressure is proportional to $\exp(-L_v/RT)$, where L_v is the latent heat of vaporization. Thus as T decreases the vapor pressure falls exponentially, and with it the pumping speed. Microscopically it is a case of having fewer and fewer energetic molecules capable of escaping to the vapor. The vapor becomes tenuous. Pumping it away carries off little heat and the rate

of cooling diminishes. A lower limit of about 0.3 K can be reached by pumping on ^{3}He, the rare isotope of helium.

An understanding of the principles of dilution refrigeration requires a familiarity with the properties of the stable helium isotopes, ^{3}He and ^{4}He, and mixtures of the two. The ^{4}He atoms are bosons; the ^{3}He atoms are fermions. This very important distinction stems from the fact that the ^{3}He nucleus contains an odd number of fermions, one neutron and two protons. The ^{4}He nucleus contains an even number of fermions, two neutrons and two protons. Both ^{3}He and ^{4}He remain liquids under their saturated vapor pressure at the lowest temperatures yet achieved. All available experimental evidence indicates they would remain liquids even at absolute zero. Solid phases can be produced by raising the pressure to near 30 atm.

^{4}He exhibits superfluid behavior below the lambda point $T_\lambda = 2.172$ K. Below 0.5 K, ^{4}He has essentially zero viscosity, zero heat capacity, zero entropy, and zero pressure. This parade of near zeroes may be traced to the momentum distribution of ^{4}He atoms below T_λ. As the temperature falls below T_λ the ^{4}He atoms begin to condense in the state of zero momentum. As bosons, there is no exclusion principle operating to prevent such a pile-up of ^{4}He atoms in the same momentum space cell. The fraction of the ^{4}He atoms in the zero momentum state increases sharply from zero, at T_λ, to one at absolute zero. In mixtures of ^{3}He and ^{4}He, the ^{4}He behaves like a mechanical vacuum; its near-zero pressure allows it to play a neutral role in so far as the mechanical aspects are concerned. The ^{4}He is essential to the dilution refrigerator for two reasons; first, it takes up space—it dilutes the ^{3}He. Of greater importance is the fact that pure ^{4}He has a greater affinity for a ^{3}He atom than does pure ^{3}He. We will touch on this point in some detail after describing the properties of ^{3}He–^{4}He mixtures.

Mixtures of ^{3}He and ^{4}He are tantalizing. Above 0.87 K the composition can be varied at will. (The ^{3}He boils away above 3.2 K.) At temperatures below 0.87 K there can be a phase separation. The mixture separates into two distinct components. The precise temperature of the phase separation depends on the relative concentrations of ^{3}He and ^{4}He. Both phases are liquids and in fact both are mixtures of ^{3}He and ^{4}He. The two phases are distinguished by the relative concentrations of ^{3}He and ^{4}He. The phase separation was discovered in 1956 by G. Walters and W. Fairbank. Figure 18.8 shows the phase diagram for ^{3}He–^{4}He mixtures. The solid line in the figure locates the phase separation boundary. The quantity X measured along the abscissa is the ^{3}He fraction

$$X \equiv \frac{N_3}{N_3 + N_4}$$

where N_3 and N_4 are the numbers of ^{3}He and ^{4}He atoms. At any temperature above 0.87 K a horizontal line on the phase diagram does not intercept the

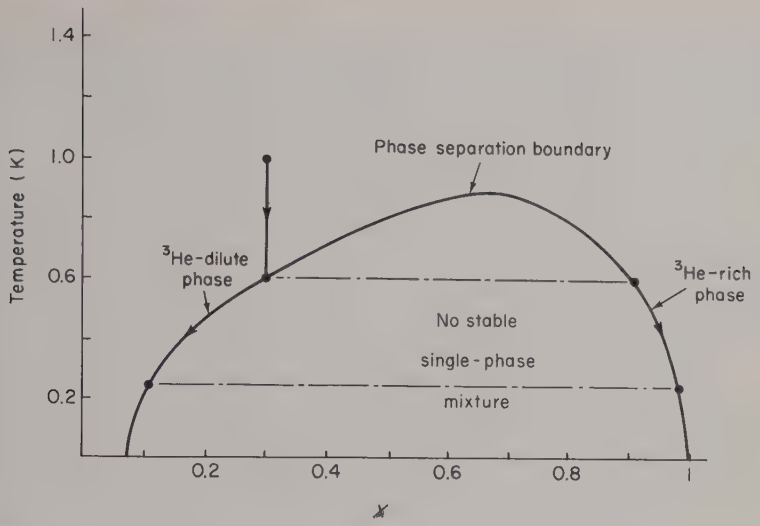


Figure 18.8. Phase diagram for ^{3}He–^{4}He mixtures, where X denotes the fraction of ^{3}He atoms in the mixture(s).

phase separation boundary, indicating that any mixing ratio of ^{3}He–^{4}He forms a stable single-phase solution. At any temperature below 0.87 K a horizontal line on the diagram intersects the phase separation boundary at two points. These two points fix the values of X for the two phases. The region beneath the phase separation boundary does not correspond to any stable thermodynamic state of ^{3}He–^{4}He mixtures.

The arrows in Figure 18.8 indicate the sequence of events which accompany the progressive cooling of a ^{3}He–^{4}He mixture. The system starts at a temperature of 1 K as a single-phase mixture in which 30% of the atoms are ^{3}He. As the mixture is cooled, nothing spectacular happens until the temperature reaches about 0.60 K. At this point there is a phase separation. The mixture separates into two distinct liquid layers. One layer is greatly enriched in ^{3}He, being nearly 90% ^{3}He. This ^{3}He-rich phase is less dense and so floats on the dilute phase. As the temperature is lowered still further, the ^{3}He-rich phase grows in size and becomes richer in ^{3}He, while the dilute phase is progressively impoverished.

Below a temperature of about 0.2 K the ^{3}He-rich phase is essentially pure ^{3}He while the dilute phase—the bottom layer—is about 6% ^{3}He and 94% ^{4}He. As the diagram reveals there will be about 6% ^{3}He in the dilute phase even at absolute zero. This unique property of ^{3}He–^{4}He mixtures was not confirmed until nearly 10 years after the discovery of the phase separation. It was expected that the ^{3}He-rich phase would behave as indicated in the diagram. The 6%

solubility of ^{3}He in the dilute phase was a welcome surprise: it is the key to dilution refrigeration. Research workers in the field felt that the ^{3}He-dilute phase would become pure ^{4}He as the temperature approached absolute zero. The arguments were based on the third law of thermodynamics and the tried and sometimes deceptive principle that "believing is seeing." The pre-experimental argument went like this: The third law decrees that the entropy tend toward zero as the temperature tends toward zero. Furthermore, a decrease in entropy means an increase in order. An obvious way to increase the order is via a phase separation, with the pure ^{3}He and pure ^{4}He forming two different layers. So went the argument before the experimental evidence argued otherwise. The argument is weak because the entropies of both pure ^{3}He and pure ^{4}He tend to zero as the temperatures approach absolute zero. The benefits accruing from a phase separation are of secondary importance. The configuration which actually prevails hinges on the question of energy. In other words, all mixing ratios have essentially the same entropy—nearly zero. It therefore becomes primarily a question of which configuration will give the lowest free energy.

We stated earlier that it is energetically favorable for ^{3}He atoms to mix with ^{4}He atoms. They are more tightly bound that way. Figure 18.9 compares answers to the questions "What energy levels are open to ^{3}He atoms added to pure ^{3}He?" and "What energy levels are open to ^{3}He atoms added to pure ^{4}He?" The figure indicates conditions at absolute zero and therefore ignores purely thermal energies. In part (a) the first ^{3}He atom added to pure ^{4}He is bound with an energy E_{34}. Subsequently added ^{3}He atoms are not as tightly bound. The Pauli exclusion principle forces the ^{3}He fermions into states of higher energy. In part (b) the first ^{3}He added to pure ^{3}He is bound with an energy E_{33}. The fact that $E_{34} > E_{33}$ means that ^{3}He will dissolve in ^{4}He rather than separating as a separate phase. There is a limit to the amount of ^{3}He which will dissolve in ^{4}He (for $T < 0.87$ K) because the Pauli principle forces each succeeding ^{3}He atom into a level of energy greater than E_{34}. Eventually it no longer becomes energetically favorable for the ^{3}He to dissolve in the ^{4}He. At temperatures below about 0.1 K the limit is near 6% ^{3}He.

It is instructive to see how theory and experiment are combined to predict the limiting value of 6% for the solubility of ^{3}He in the dilute phase. Thermodynamically, phase equilibrium requires equality of the chemical potentials of the two phases. From a microscopic viewpoint phase equilibrium results when a ^{3}He atom does not find it energetically more favorable to join one phase than the other. This interpretation permits us to write the condition for phase equilibrium as

energy of ^{3}He atom added to the ^{3}He-rich phase
= energy of ^{3}He atom added to dilute solution of ^{3}He–^{4}He

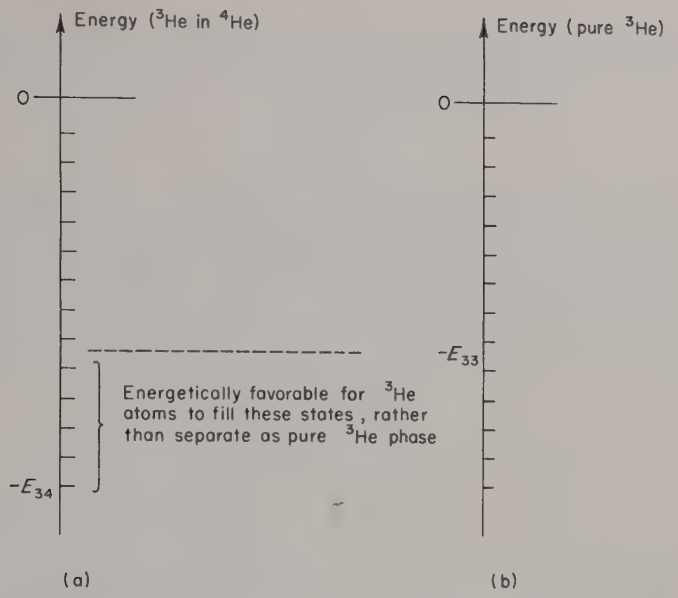


Figure 18.9. Schematic of the energy levels open to (a) ^{3}He atoms in a ^{3}He–^{4}He mixture and in (b) pure ^{3}He at absolute zero. A single ^{3}He atom is bound more tightly in ^{4}He than in pure ^{3}He. The Pauli exclusion principle forces additional ^{3}He atoms into states of higher energy, but these states are of lower energy than the states open in pure ^{3}He. Thus ^{3}He will dissolve in ^{4}He rather than separate as a second phase.

This is expressed symbolically as

$$\mu_{33} = \mu_{34}$$

Because we restrict ourselves to temperatures near absolute zero where the purely thermal energies are ignorable, expressions for μ_{33} and μ_{34} follow at once from Figure 18.9. At sufficiently low temperatures the ^{3}He-rich phase is nearly pure ^{3}He. This fact allows us to write

$$\mu_{33} \simeq -E_{33} \tag{18.47}$$

where E_{33} is the energy required to remove a ^{3}He atom from pure ^{3}He. Experimentally, E_{33} is the latent heat (per atom) of pure ^{3}He. Figure 18.9 reveals the energy of a ^{3}He atom in the dilute phase as a sum of kinetic and potential energies

$$\mu_{34} = -E_{34} + \varepsilon_F \tag{18.48}$$

The negative term, $-E_{34}$, is the potential energy with which the ^{3}He atom is

bound to its neighbors in the liquid. The positive term, ε_F, is the kinetic energy forced upon the atom by the Pauli exclusion principle, and this equals the Fermi energy of the added ^{3}He atom [compare Eq. (17.16)]

$$\varepsilon_F = \frac{\hbar^2}{2m}(3\pi^2 n_3)^{2/3} \tag{18.49}$$

The mass m in (18.49) is not the mass of a ^{3}He atom. The effective mass of the ^{3}He is increased by virtue of the interaction with neighboring atoms. Experiments which measure the speed at which temperature differences propagate through ^{3}He–^{4}He mixtures show that m is equal to 2.5 times the mass of a ^{3}He atom. For a dilute mixture the ^{3}He fraction is related to the number density of the ^{3}He by

$$X \equiv \frac{N_3}{N_4+N_3} \simeq \frac{N_3}{N_4} = \frac{n_3}{n_4}$$

where n_4 is the number density of ^{4}He. Thus the Fermi energy term is proportional to $X^{2/3}$

$$\varepsilon_F = \frac{\hbar^2}{5m_3}(3\pi^2 n_4)^{2/3} X^{2/3} \tag{18.50}$$

With these results the condition for phase equilibrium becomes

$$-E_{33} = -E_{34} + \frac{\hbar^2}{5m_3}(3\pi^2 n_4)^{2/3} X^{2/3} \tag{18.47a}$$

The energy E_{34} depends on X. This dependence is determined from measurements of the temperature at which the phase separation occurs. The experiments conducted by Edwards and his co-workers show that $E_{34}(X) - E_{33}$ increases linearly with X. Over the experimental range of X their results may be represented by

$$E_{34}(X) - E_{33} \simeq 2.7(1+14X) \times 10^{-17} \text{ erg}$$

The values of n_4 and m_3 are

$$n_4 \simeq 2.2 \times 10^{22} \text{ cm}^{-3}, \qquad m_3 \simeq 5 \times 10^{-24} \text{ gm}$$

The limiting value of X emerges as the solution of

$$1 + 14X = 12X^{2/3} \tag{18.51}$$

As the student may verify, the solution is $X \simeq 0.06$; at absolute zero the dilute phase will contain about 6% ^{3}He.

.....

18.4.1 Verify that the solution of (18.51) is $X \simeq 0.06$ by locating the intersection of the graphs of $1 + 14X$ and $12X^{2/3}$.

.....

Earlier in this section we noted that the adiabatic expansion of a gas is one of the classic methods of cooling. The analogous way of cooling ^{3}He is by diluting it with ^{4}He. At temperatures below 0.5 K the ^{4}He has a negligible heat capacity. Dilution therefore enlarges the volume occupied by the ^{3}He without any appreciable heat transfer, that is, it expands the ^{3}He adiabatically. A dilution refrigerator based on this type of adiabatic expansion has been operated by deBruyn and his colleagues.† This type of refrigerator is suited only to single-cycle operation; it is a one-shot process. We turn next to the description of a dilution refrigerator which may be operated continuously. This second version of a dilution refrigerator is analogous to the forced evaporation type of classic cooler.

The cooling occurs in the so-called mixing chamber, where a layer of nearly pure ^{3}He floats on the ^{3}He-dilute mixture. At 0.2 K, for example, the upper layer is about 99% ^{3}He; the lower layer is roughly 8% ^{3}He. The ^{3}He is withdrawn from the dilute solution, circulated through heat exchangers, and so on, and eventually returned to the upper layer. As ^{3}He is removed from the bottom layer, ^{3}He atoms move downward from the upper layer in an effort to maintain the phase equilibrium. In moving across the phase boundary the ^{3}He carries away latent heat. This heat eventually leaves the mixing chamber by way of the ^{3}He circulation. This mechanism operates just like forced evaporation. The upper layer of ^{3}He acts like the liquid being evaporated. The ^{3}He in the dilute lower layer plays the role of a vapor. It is pumped away and carries with it the latent heat. So, if you take a schematic diagram of an evaporation cooler, turn it upside down, label the liquid "pure ^{3}He," and label the vapor "dilute ^{3}He–^{4}He mixture," you end up with a schematic of the mixing chamber segment of a dilution refrigerator.

There is a strong analogy between latent heats as well. In both cases the atoms which move across the phase boundary are those with much higher than average energy. The hot atoms are pumped away, leaving the cooler ones behind. One of the virtues of the dilution refrigerator is that the dilute ^{3}He phase—which corresponds to the vapor—never falls below a concentration of 6%. By contrast, the vapor pressure of the classic cooler vanishes exponentially as T tends toward absolute zero. In short, ^{3}He is self-refrigerating; classic coolers are not.

---

† A schematic diagram of this device is shown as Fig. 19 of the article by J. C. Wheatley, Dilute Solutions of ^{3}He in ^{4}He at Low Temperatures, *Am. J. Phys.* **36**, 181 (1968).

18.5 SUPERCONDUCTIVITY

In 1908 Heike Kammerlingh Onnes succeeded in liquifying helium. Three years later he discovered that mercury loses all electrical resistance below a temperature of 4.17 K. This phenomenon is called *superconductivity*. Since the initial discovery by Onnes, several thousand compounds and over one-third of the elements have been found to exhibit superconductivity. The behavior of the electrical resistivity ρ as a function of temperature is shown in Figure 18.10

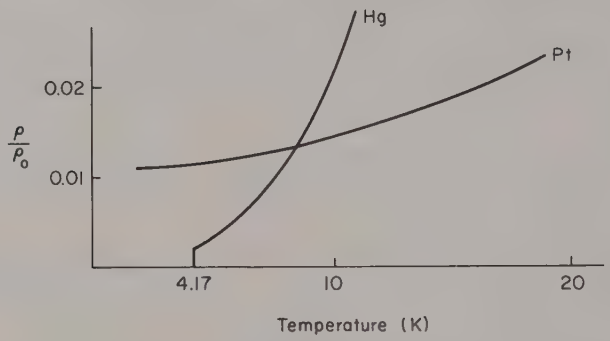


Figure 18.10. Resistivity versus temperature for mercury and platinum. The quantity ρ_0 is the resistivity at 0 °C. Mercury becomes superconducting at its transition temperature of 4.17 K.

for mercury and for platinum, a nonsuperconductor. The quantity ρ_0 denotes the resistivity at 0 °C. Platinum was one of the first materials investigated by Onnes. He attributed the leveling off of ρ/ρ_0 to impurities. He then turned to mercury which can be purified by vacuum distillation. Onnes expected that the resistivity of a pure metal would fall toward zero as the temperature approached absolute zero. The abrupt[†] disappearance of electrical resistance at 4.17 K was completely unexpected. Onnes subsequently found that tin and lead also become superconducting at sufficiently low temperatures. The resistivity of a good conductor at 0 °C is on the order of 10^{-5} ohm-cm. At liquid helium temperatures it has dropped by a factor of 100 or so to, say, 10^{-7} ohm-cm. Onnes showed that the resistivity of superconducting lead is less than 10^{-16} ohm-cm by observing the persistent current induced in a lead ring. The Onnes ring experiment consists of three steps:

(a) A lead ring is placed in the field of a magnet and then cooled below its superconducting transition temperature by bathing it in liquid helium.

[†] The transition from the normal to the superconducting state takes place over an interval of less than 0.1 K.

(b) The magnet is removed, inducing a "super" current in the ring. Since the current experiences no resistance it persists indefinitely.

(c) The persistent current induced in the ring creates a magnetic field which may be detected, for example, by observing the deflection of a compass needle. Observations over a period of a few hours established that the current did not diminish by as much as 1%, the precision of the measurement. Recent observations of persistent currents over a period of months have established $\rho < 10^{-26}$ ohm-cm.

The experimental data are consistent with the presumption that $\rho = 0$ for a superconductor. This means that there can be no electric field E inside a superconductor. This conclusion follows from Ohm's law which specifies the relation between E, ρ, and the electric current density J,

$$E = \rho J$$

Since $\rho = 0$ and J is finite, E must vanish.†

The temperature at which a material becomes superconducting is termed the *transition temperature*. The value of the transition temperature depends upon whether or not an external magnetic field is present. Transition temperatures for several superconductors are listed in Table 18.1. The temperatures are zero magnetic field values. The effects of an applied magnetic field will be discussed shortly.

TABLE 18.1

Transition Temperatures for Type I and Type II Superconductors

Type I	T_0 (K)	Type II	T_0 (K)
Lead (Pb)	7.19	Nb_3Sn	18.05
Mercury (Hg)	4.17	V_3Si	17.0
Tin (Sn)	3.72	Niobium (Nb)	8.70
Indium (In)	3.40	Vanadium (V)	4.89
Aluminum (Al)	1.20	Tantalum (Ta)	4.38

There appear to be two distinct types of superconducting materials. Type I, or soft superconductors include elements like Hg, Pb, Bi, In, Al, and Sn: they are mechanically soft and malleable, thus the terminology. By contrast, type II superconductors which include niobium (Nb, formerly called Columbium) and tantalum are hard and brittle, a troublesome property when trying to draw such materials into fine wires. The type II superconductors are a breed by

† Strictly speaking these conclusions are only correct for direct currents and static fields. In very-high-frequency fields the electrical resistivity of a superconductor does not vanish.

themselves and only in recent years have they been subjected to intensive study. The reasons for the great interest in type II superconductors will be discussed later.

In addition to the disappearance of electrical resistance, type I superconductors display several other distinctive features. We consider five of these. In the course of the discussion we develop the thermodynamics of superconductivity and comment on some aspects of the microscopic theory of superconductivity. The section concludes with remarks about type II superconductors and the quantization of magnetic flux.

A Meissner Effect

In 1933 Meissner and Ochsenfeld performed an experiment which revealed that a superconductor is not only a perfect conductor, but a perfect diamagnet as well. They observed the following: When a tin cylinder was (a) cooled below its superconducting transition temperature and then (b) subjected to a magnetic field, the lines of magnetic flux did not penetrate the sample. This is the behavior expected for a perfect conductor. As the magnetic field is turned on currents are induced tending to prevent any change in the magnetic flux through the sample (Lenz' Law). In the absence of any resistance this current assumes whatever value is necessary to prevent flux penetration.

The unexpected occurred when Meissner and Ochsenfeld reversed the order of (a) and (b). The magnetic field was first applied to the sample in its normal, that is, nonsuperconducting, state. The magnetic properties of the normal state are very weak, so that the magnetic flux lines fully penetrated the sample. The material was then cooled below its transition temperature. If zero resistivity were the only attribute of superconductivity, the flux threading the specimen would have remained unchanged as the transition temperature was passed. What Meissner and Ochsenfeld observed was something quite different. They found that the magnetic flux was completely expelled from the sample as it became superconducting. This perfect diamagnetism means that the magnetic induction B inside a superconductor is zero.[†] Thus $E = 0$ for a perfect conductor and $B = 0$ for a perfect diamagnet. The superconducting state exhibits both of these extreme properties.

The importance of the perfect diamagnetic behavior of a superconductor cannot be overemphasized. If superconductivity were solely a matter of zero resistivity no thermodynamic treatment of the phenomenon would be possible. Thermodynamically, the normal $\leftrightarrow$ superconducting transition qualifies as a

[†] As we point out in Section C an applied magnetic field does penetrate the surface skin layers to a depth of about 10^{-5} cm.

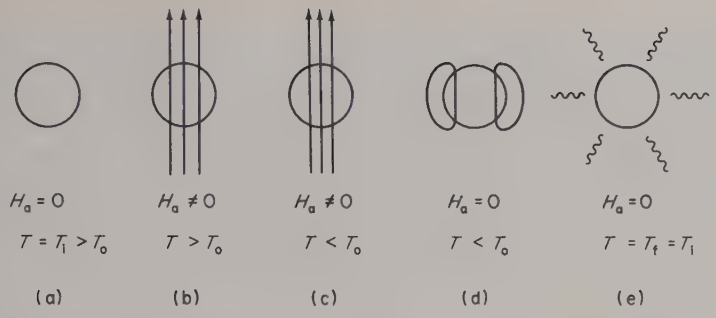


Figure 18.11. (a) If a superconductor were only a perfect conductor the $s \rightarrow n$ transition would be irreversible ($H_a = 0$, $T = T_i > T_0$). (b) The field is applied ($H_a \neq 0$, $T > T_0$) and (c) the sample cooled below its transition temperature ($H_a \neq 0$, $T < T_0$) at which point flux lines still thread the sample. (d) When the external field is removed flux is trapped, storing magnetic energy in the material ($H_a = 0$, $T < T_0$). (e) This energy is transferred to the surroundings as heat in returning the sample to its initial state ($H_a = 0$, $T = T_f = T_i$). The net result is the conversion of magnetic energy into heat—an irreversible action.

reversible process. In fact, it is a phase transition. However, if a superconductor were not a perfect diamagnet, the transition would be an irreversible process.

The concept of a reversible process was developed at length in Chapter 4. Figure 18.11a–e illustrates how irreversibility would arise if a superconductor exhibited zero resistivity but were not a perfect diamagnet. Figures 18.12a–e show how perfect conductivity and diamagnetism combine to achieve reversibility in a superconductor. The key point is that a perfect conductor would trap

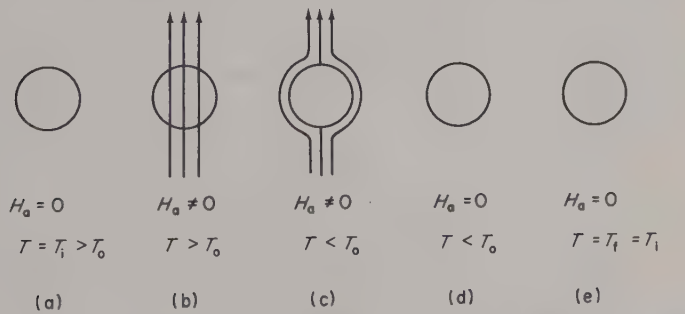


Figure 18.12. (a) The sequence of operations depicted in Figure 18.11 results in a reversible process for an actual superconductor ($H_a = 0$, $T = T_i > T_0$). (b) Flux penetrates the sample when the temperature exceeds the transition temperature T_0 ($H_a \neq 0$, $T > T_0$). (c) When the sample is cooled below T_0, flux is expelled ($H_a \neq 0$, $T < T_0$). (d, e) The system returns reversibly to its initial state because no magnetic energy has been stored in the superconducting material [(d) $H_a = 0$, $T < T_0$; (e) $H_a = 0$, $T = T_f = T_i$)].

flux (Figure 18.11d) thereby storing magnetic energy. This energy would be converted into heat when the material returned to the normal state (Figure 18.11e). The net result of the process would be the conversion of magnetic energy into heat—an irreversible action. By contrast, an actual superconductor expels flux. Magnetic energy is stored in the external field, not in the superconductor. There is no flux trapping and hence no thermalization of magnetic energy when the material returns to its normal state. We remarked earlier that the reversibility of the superconducting transition makes possible a thermodynamic treatment. The thermodynamics of superconductivity is developed in Section D.

B Critical Magnetic Field

In 1914 Onnes discovered that a superconductor reverts to its normal state if subjected to a magnetic field of sufficient strength. The critical field strength at which superconductivity is quenched depends on the temperature. Experimentally the relation between the critical field and temperature is closely approximated by the parabolic form

$$H_c = H_0\left[1 - \left(\frac{T}{T_0}\right)^2\right] \tag{18.52}$$

H_c denotes the value of the applied field strength for which the superconducting $\leftrightarrow$ normal transition occurs at the temperature T. The zero-field transition temperature is T_0 and H_0 is the critical field at absolute zero. The values of T_0 and H_0 are characteristics of the particular superconductor. Table 18.2 lists a few representative values. A plot of the relation between H_c and T, Figure 18.13, amounts to a phase diagram for a superconductor. The phase boundary defined by (18.52) divides the H–T plane into two regions—the normal and superconducting phases.

TABLE 18.2

Critical Fields and Transition Temperatures for Type I Superconductors

Material	H_0 (G)	T_0 (K)
Al	100	1.20
In	275	3.40
Sn	300	3.72
Hg	420	4.17
Pb	800	7.19

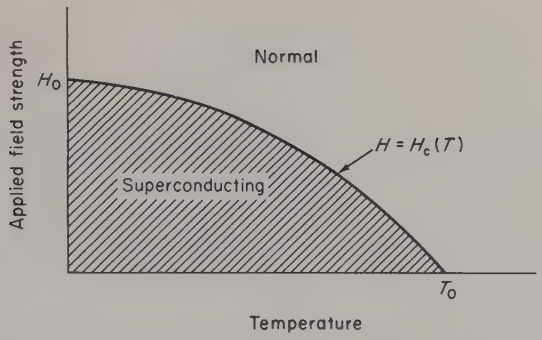


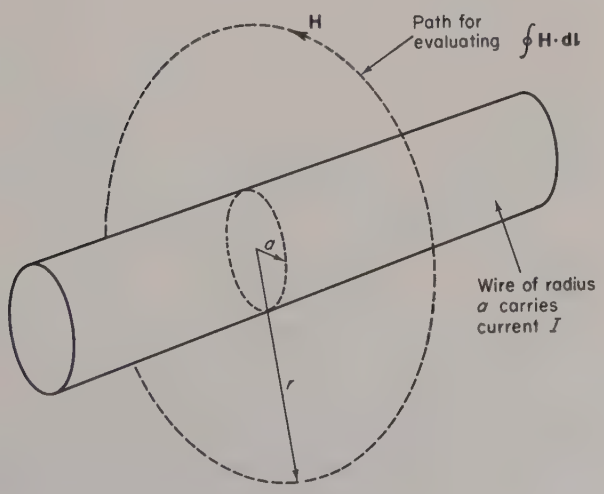
Figure 18.13. The phase boundary $H = H_c(T)$ divides the H–T plane into two regions. Experimentally the equation defining the boundary has very nearly the parabolic form of Eq. (18.52).

It was also observed that superconductivity vanished whenever the current carried by a superconductor exceeded some critical value (dependent on material)—even in the absence of an external magnetic field set up by the current itself. Experiment showed this to be the case. Recall Ampere's Law

$$\oint \mathbf{H}\cdot\mathbf{dl} = \frac{4\pi}{c} I_{\text{enc}}$$

Referring to the accompanying diagram shows

$$H(r)2\pi r = \frac{4\pi}{c} I, \qquad r > a$$



For $r = a$ we obtain H at the surface of the wire

$$H_{sur} = \frac{2I}{ca} \tag{18.53}$$

When the current I is large enough, H_{sur} exceeds the critical field and superconductivity is quenched. This current is referred to as the *critical current.* When Onnes first discovered superconductivity he realized at once its great potential for producing strong magnetic fields without having to pay for the Joule heating. His disappointment was equally keen when he found that the superconducting current produced a self-destructing magnetic field. The recent exploitation of type II superconductors has seen Onnes' dream turn into reality.

C Penetration Depth

A superconductor does not completely expel a magnetic field. Experiments show that an applied field penetrates the surface, but that the field decreases exponentially with depth (see Figure 18.14). The characteristic penetration depth λ is about 10^{-5} cm. Thus any applied magnetic field is confined to the surface skin layers of a superconductor. It follows from Ampere's Law that any supercurrents must flow in the skin of the material, that is, from

$$H(r) = \frac{2I_{enc}}{cr}$$

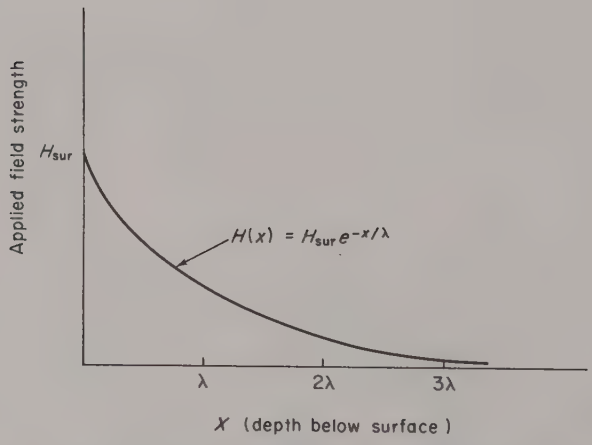


Figure 18.14. An applied magnetic field decreases exponentially with depth below the surface of a material in the superconducting state. The penetration depth λ is in the 10^{-6}–10^{-5} cm range.

it follows that, if H decreases exponentially beneath the surface, so also must the enclosed current. In passing we note that there is no corresponding penetration depth for a static electric field. This was demonstrated by H. London who measured the capacitance of a capacitor having superconducting plates.

D Phase Transition

Thermodynamically, the normal ↔ superconducting transition qualifies as a phase change. In the presence of an applied magnetic field a latent heat is rejected as a substance goes from the normal to superconducting phase. There is no latent heat in the absence of a magnetic field. Further, the specific heat of a superconductor undergoes a discontinuous change at the transition temperature (see Figure 18.15). The details of these effects follow from the thermodynamics of superconductivity which we now develop.

In Chapter 4 we showed that the magnetic work done by a system could be expressed in the form

$$dW = -H\,dM \tag{4.15}$$

The total thermodynamic work is the sum of the mechanical work and the magnetic work

$$dW = P\,dV - H\,dM \tag{18.54}$$

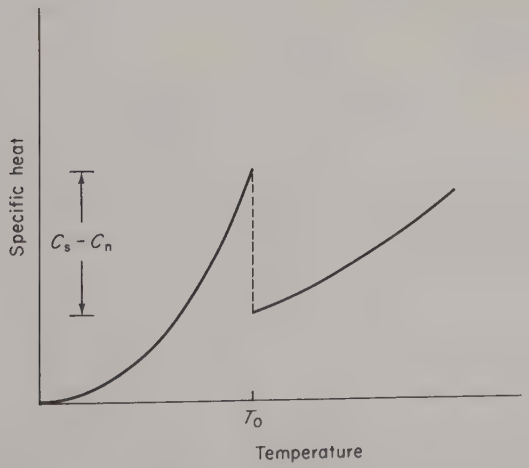


Figure 18.15. A superconductor exhibits a specific heat discontinuity at the transition temperature T_0.

The mechanical work is generally negligible compared to the magnetic work.† It is helpful to carry both terms because the analogs

$$P \leftrightarrow H, \qquad dV \leftrightarrow -dM \tag{18.55}$$

permit an immediate generalization of the equations for the thermodynamic potentials and their differentials. For example, consider the Gibbs free energy G

$$G(T,P) = U - TS + PV \tag{8.12'}$$

and its differential

$$dG = -S\,dT + V\,dP \tag{8.13'}$$

An appeal to (18.55) shows that we can include magnetic effects by writing

$$dG = -S\,dT + V\,dP - M\,dH \tag{18.56}$$

By regarding G as a function of the independent variables T, P, and H we obtain the generalization of (8.14)

$$\begin{aligned} S &= -\left(\frac{\partial G}{\partial T}\right)_{P,H} \\ V &= \left(\frac{\partial G}{\partial P}\right)_{T,H} \\ M &= -\left(\frac{\partial G}{\partial H}\right)_{T,P} \end{aligned} \tag{18.57}$$

The Gibbs free energy is

$$G = U - TS + PV - \int_0^H M(H)\,dH \tag{18.58}$$

This form of G ignores any H dependence of U and S. The first three terms on the right side are essentially independent of H. The magnetic field effects are incorporated in the last term. Ordinarily, the magnetization M is related to the magnetic induction B and the applied magnetic field H by

$$B = H + \frac{4\pi M}{V} \tag{4.11}$$

---

† Moderate changes in pressure have little effect on the properties of superconducting materials. However, a number of elements (for example, Te, Ge, Se), which are not superconducting at ordinary pressures, do make the transition at pressures of many thousands of atmospheres.

If we apply this to a superconductor wherein $B = 0$, we have

$$M = -\frac{VH}{4\pi} \tag{18.59}$$

The last term in (18.58) then gives

$$-\int_0^H M(H)\,dH = \frac{V}{4\pi}\int_0^H H\,dH = \frac{VH^2}{8\pi} \tag{18.60}$$

With this result we can write the Gibbs free energy of the superconducting phase as

$$G_s(T,P,H) = G_s(T,P,0) + \frac{VH^2}{8\pi} \tag{18.61}$$

where $G_s(T,P,0)$ comprises the field-independent terms of (18.58). The normal phase of a superconductor displays only a very weak dependence on a magnetic field. It is so weak in fact, that we can ignore the magnetic contribution to $G_n(T,P)$, the Gibbs free energy of the normal phase.

Equilibrium between the normal and superconducting phases—described pictorially in Figure 18.13—requires equality of the Gibbs free energies of the normal and superconducting phases. We ignore the pressure as a variable and concentrate our attention on the H, T dependence. The phase equilibrium condition reads

$$G_n(T) = G_s(T,0) + \frac{VH_c^2}{8\pi} \tag{18.62}$$

This equation is the thermodynamic relation defining the n–s phase boundary in Figure 18.13.

The relationship between the entropies of the normal and superconducting phases follows from

$$S = -\left(\frac{\partial G}{\partial T}\right)_{P,H} \tag{18.57a}$$

namely,†

$$S_n = S_s - \frac{VH_c}{4\pi}\frac{dH_c}{dT} \tag{18.63}$$

Experimentally, dH_c/dT is negative for all superconductors. It follows that $S_s < S_n$; the entropy of the superconducting phase is less than that of the

† Note the distinction between H the applied field and H_c the field which defines the phase boundary. The former is an independent thermodynamic variable. The latter is a temperature-dependent quantity [compare Eq. (18.52)].

normal phase. This in turn implies that the superconducting state is one of greater order than the normal state. The origin of the increased order in the superconducting phase lies in an attractive interaction between electrons. The nature of this attractive electron–electron interaction is discussed later in our outline of the theory of superconductivity.

The latent heat absorbed by a specimen going from the superconducting to normal phase is

$$L_s = T(S_n - S_s) = -T\frac{VH_c}{4\pi}\frac{dH_c}{dT} \tag{18.64}$$

Note that the latent heat vanishes in the absence of an external field. The observed values of latent heats are in complete accord with this result, thereby confirming the reversible character of the transition.

....

18.5.1 In (18.63) we showed that $S_s < S_n$. Stability of an equilibrium state also requires that $F_s < F_n$, where F denotes the Helmholtz free energy. Make use of the analogs (18.55) to obtain an expression for F which includes a magnetic term. Show that the difference of F_n and F_s at the phase boundary is given by

$$F_n - F_s = \frac{VH_c^2}{8\pi}$$

....

By combining $dQ = T\,dS$ and the definition of the specific heat $dQ = C\,dT$ we find

$$C = T\left(\frac{dS}{dT}\right)$$

Applying this to (18.63) we obtain the specific heat† discontinuity

$$C_s - C_n = \frac{TV}{4\pi}\left[H_c\left(\frac{d^2H_c}{dT^2}\right) + \left(\frac{dH_c}{dT}\right)^2\right] \tag{18.65}$$

Unlike the latent heat, the discontinuity in the specific heat is nonzero even in the absence of an applied field. Experimentally, the discontinuity is about twice C_n in the absence of an applied field ($H_c = 0$). The heat capacities become equal when $H_c \simeq \frac{2}{3}H_0$. In stronger fields C_n exceeds C_s (see Problem 18.5.2.)

† Equation (18.65) gives the total heat capacity of a sample of volume V. To obtain the specific heat capacity it is necessary only to replace V by the reciprocal of the mass density.

.....

18.5.2 (a) Prove that L_s is positive (that is, heat is absorbed in the s → n transition) for the parabolic form of $H_c(T)$ given by (18.52).

(b) Using (18.52) show that

$$C_s - C_n = \frac{VH_0^2}{2\pi T_0} x(3x^2 - 1)$$

where $x = T/T_0$.

(c) Show that $C_s = C_n$ when $H_c = \frac{2}{3}H_0$.

(d) Demonstrate that a superconductor which follows (18.52) obeys the third law of thermodynamics. [*Hint:* Show that the entropy difference $S_s - S_n$ vanishes at absolute zero.]

.....

Before moving on to consider the fifth and final notable property of superconductors, a pedagogic confession is in order. It may be misleading to infer from $B = 0$ that $M = -VH/4\pi$, as was done earlier. The implication of this relation is that $H \neq 0$ even though $B = 0$. The actual state inside a superconductor has $M = 0$, $H = 0$, as well as $B = 0$. The two descriptions turn out to be equivalent because the former description ignores the induced super currents. These currents flow in the skin and shield the interior from the applied field. The advantage of using $M = -VH/4\pi$ is that it allows us to employ the conventional expression for magnetic work $-H\,dM$. This in turn leads to the magnetic version of the Gibbs free energy, the condition for phase equilibrium, and so on. The student may rest assured that the same results follow—if less swiftly—from a more proper analysis.

E Isotope Effect

The transition temperature T_0 is not the same for different isotopes of certain superconducting elements. In 1950 Serin and his co-workers and, independently, Maxwell, observed the zero-field transition temperature of several isotopes of mercury. They found that the transition temperature decreased with increasing mass number M. Their measurements indicated that†

$$M^{1/2}T_0 = \text{constant}$$

In one of the few instances in the brief history of superconductivity where theory preceded experiment, Fröhlich (1950) suggested a mechanism which

† Experimentally, Serin and co-workers found $M^{\alpha}T_0 = \text{constant}$, with $\alpha = 0.504$; Maxwell found $\alpha = 0.486$. The theory advanced by Fröhlich implied $\alpha = \frac{1}{2}$. Subsequent studies revealed superconducting elements with values of α differing significantly from $\frac{1}{2}$.

might give rise to superconductivity. His was not a complete theory of superconductivity but it did predict the isotope effect.

Our presentation of the theory of superconductivity is necessarily superficial. Quantum mechanics is indispensible in any treatment which would furnish a deep understanding of superconductivity. Considering that the development of a satisfactory microscopic theory defied a generation of theoretical physicists, we offer no apologies for our abbreviated remarks.

At first thought it seems that the occurrence of superconductivity is a miracle. This impression is prompted by a calculation of the energy difference between the normal and superconducting phases. This difference is extremely small—on the order of 10^{-7} eV/atom. The figure of 10^{-7} eV/atom can be argued as follows: The energy difference between the normal and superconducting phases is of order $VH_0^2/8\pi$—the energy increase for a superconducting-to-normal transition near absolute zero. The energy difference per atom is $H_0^2/8\pi n$, where n is the number of atoms per unit volume. Taking $H_0 = 300$ G and $n \simeq 10^{22}$ cm^{-3} leads to the 10^{-7} eV/atom value.

....

18.5.3 Verify the estimate of 10^{-7} eV/atom for mercury and aluminum. Consult Table 18.2 for H_0 values.

....

By contrast the binding energy which glues the atoms into a lattice is on the order of a few electron volts per atom. Similarly, at 1 K the random thermal energy of vibration associated with the atomic lattice is on the order of 10^{-4} eV/atom. Thus it would seem that superconductivity is almost a miracle—requiring a delicate balance between several big effects before it can be observed. In a sense this is the case—there are many effects which cancel because they are unchanged by the transition. For example, the metallic lattice of atoms and ions is not changed by the superconducting transition. This fact was established long ago by Keesom and Onnes through X-ray diffraction. The particles directly responsible for the superconducting transition are the valence electrons. These are the electrons which, in a normal metallic state, are more or less free to move in response to an external electric field. However, the valence electrons interact with each other; they interact with defects in the lattice, with vibrations of the lattice, and with impurities randomly distributed throughout the material. It is necessary to isolate the interaction which precipitates the superconducting transition and ignore the rest—on the grounds that these are the same for both phases. Singling out the essential interaction was a crucial step in the development of a microscopic theory of superconductivity.

It was here that Fröhlich's work and the experimental confirmation of the isotope effect attained their great significance. The interaction giving rise to the

isotope effect is called the *electron–phonon interaction.* It gives rise to an *attractive* force between electrons. This attraction fosters the increased order, that is, the lower entropy, characteristic of the superconducting phase. The attractive interaction springs from a polarization of the lattice. Each electron distorts the lattice in its immediate vicinity. The distortion of the lattice at one point is felt throughout the solid—the disturbance propagating as quantized sound waves, or phonons, as they are called. Other electrons feel the effects of such a distortion because the configuration of the lattice in their immediate vicinity is altered by the phonons. Thus the electrons interact with each other via the phonons.

The attractive electron–electron interaction, mediated by the phonons, is opposed by the repulsive Coulomb force. The effective Coulomb energy between electrons separated by a distance r is not e^2/r because of *shielding.* A shielding of the Coulomb interaction results because, on the average, there is a net positive charge in the vicinity of an electron and a net negative charge surrounding a positive ion. Each charge continually strives to surround itself with particles of opposite charge. The result may be described as one in which each charge is dressed in a screening cloud of charge of the opposite sign. The potential energy of two dressed electrons is reduced from its Coulomb value by an exponential factor $\exp(-r/r_s)$. The quantity r_s is a screening length, roughly equal to the mean distance between electrons. Thus two electrons separated by a distance of a few times r_s repel each other rather weakly. Dynamically they are hidden from each other by their screening clouds. The screening of the repulsive Coulomb forces makes it possible for the attractive electron–phonon interaction to compete effectively and give a net attraction between two electrons.

In 1956 L. N. Cooper showed how an attractive electron–electron interaction would promote the formation of electron pairs. The energetically most favorable configuration of such a pair is one in which the two electrons have antiparallel spins and equal but opposite linear momenta. Thus the combined spin angular momentum of the pair is zero, as is its total linear momentum. When a substance becomes superconducting, the attractive interaction between electrons causes a condensation of these Cooper pairs, as they are now called. At absolute zero all of the electrons form Cooper pairs. The number of Cooper pairs decreases with increasing temperature, falling to zero at the transition point.

The attractive electron–electron interaction is intrinsically weak, but has a long range—roughly 10^{-4} cm. This is more than a thousand times greater than the typical screening length r_s which defines the range of the repulsive Coulomb force. The range of the attractive electron–electron interaction is usually referred to as the *coherence length.* The coherence length may also be thought of as the characteristic dimension of a Cooper pair.

Somewhat less energetically favorable is the formation of Cooper pairs which have zero spin angular momentum but nonzero linear momentum. Such pairs travel without resistance—their motions produce the super current. In a current-carrying superconductor the full assembly of Cooper pairs is locked in step. Their collective motions result in a quantum system of macroscopic size. The only way to introduce resistance is to disrupt the entire ensemble of pairs, that is, one must destroy the superconducting state.†

In 1957 a microscopic theory of superconductivity was formulated by Bardeen, Cooper, and Schrieffer. The BCS theory assumes that the superconducting state forms as a strongly correlated assemblage of Cooper pairs. Bardeen, Cooper, and Schrieffer showed that such a state is energetically favored over one formed from normal electrons whenever the attractive electron–phonon interaction dominates the repulsive Coulomb force between electrons. The BCS theory exhibits a Meissner effect, the phase change and specific heat discontinuity, the isotope effect, and the other electrodynamic properties of superconductors. Furthermore, the theory predicted new effects which have since been verified experimentally. It seems safe to say that although the theory of (type I) superconductivity has not achieved its final form, the mechanism by which the phenomenon operates has been identified.

In view of the successes of the BCS theory this last remark may seem too modest a claim. There remain, however, many aspects of superconductivity which are not understood. For example, a superconductor with a transition temperature of 25 K–30 K would revolutionize the technologies of power transmission and communication. At present, an alloy of Nb, Al, and Ge has the highest transition temperature, 21 K. The search for these relatively high-temperature superconductors has been sparked by B. T. Matthias. Despite its many virtues, the BCS theory is powerless to suggest how one might engineer higher transition temperatures. Matthias has discovered empirical rules which suggest combinations of elements likely to have a high transition temperature. Unfortunately, many of the most promising compounds are chemically unstable. Type II superconductivity is another area where the BCS theory is not directly applicable. The distinctive characteristic of type II superconductors is the absence of a sharp Meissner effect. Magnetic flux is excluded until a critical field H_{c1} is reached. When the magnetic field is increased beyond H_{c1}, there is a gradual penetration of flux. There is a second critical field H_{c2}, greater than H_{c1}, at which flux penetration becomes complete. Above H_{c2} no

† Again we remark that this is strictly true only for direct current. At any temperature above absolute zero there will be some normal electrons. In dc operation the super current short circuits the normal electron flow. In a high-frequency electric field the inertia of the super electrons limits their current, enabling a finite resistive current of normal electrons to develop.

vestige of superconductivity remains. The range of field strengths defined by

$$H_{c1} < H < H_{c2}$$

is called the *mixed state* since portions of the sample are superconducting while others are normal. The flux penetration takes the form of flux lines which lead magnetic field lines through the sample. The flux lines arrange themselves in a triangular lattice. For fields only slightly greater than H_{c1} most of the sample is in the superconducting state. On and near the flux lines the material is normal. As the applied field strength is raised the density of flux lines increases. At the upper critical field H_{c2} the density of flux lines reaches a point where the surrounding normal regions overlap, leaving no superconducting paths through the material.

The upper critical field is generally much larger than H_{c1} or typical H_0 values for type I materials. For Nb_3Sn, $H_{c2} = 88{,}000$ G, and materials with H_{c2} as large as 400,000 G are known. Other high-field superconductors will doubtless be discovered. High-field superconductors have been used to fabricate extremely powerful electromagnets. Power is consumed only in the cryogenic system which maintains the coils in the superconducting state.

The BCS theory is not sufficiently general to deal with the mixed state of type II superconductors. Historically, a phenomenological theory encompassing the mixed state was developed by Landau and Ginzburg in 1950. The Landau–Ginzburg equations were solved by the Russian physicist A. A. Abrikosov in 1957. Abrikosov's solution predicted the triangular array of flux lines characterizing the mixed state. Furthermore, his calculations indicated that the flux which threaded the specimen was quantized. Each flux line was associated with a single flux quantum. The quantum of flux is the *fluxoid*

$$\Phi = \frac{hc}{2e} = 2.07 \times 10^{-7} \text{ G-cm}^2 \tag{18.66}$$

The existence of such flux quanta was verified experimentally in 1961 by W. Fairbank and Deaver and by Doll and Nabauer. More recently, the triangular flux line pattern has been studied directly using a novel photographic technique (see References).

The world of superconductivity shivers in the shadow of absolute zero, the unattainable minimum for temperatures. In the next and final section we discover that there is also a maximum temperature—the ultimate temperature.

18.6 THE ULTIMATE TEMPERATURE

Traditionally a textbook is not a proper place for speculative ideas. Nevertheless, in this final section of this final chapter we wish to bring the student to one of the many places along the frontiers of physical science.

Current research unites two areas, seemingly poles apart. These are the fields of high-energy physics and thermodynamics. The partnership has led to an exciting discovery: There is an *ultimate temperature*—an upper limit for temperature which cannot be exceeded regardless of how much energy is poured into a system. The value of this ultimate temperature is slightly over 10^{12} K. This temperature is established by fitting data from high energy scattering experiments to a theoretical model. Later in this section we show how the ultimate temperature may be deduced from theoretical considerations.

The origin of an ultimate temperature can be made plausible by consideration of the following thought experiment: Take 1 mole of molecular hydrogen confined in a fixed volume and add energy. Keep adding energy and note the temperature at which significant events occur. Further, note the behavior of the heat capacity C_v. A running record of the experiment might look something like that in the accompanying tabulation.

Temperature (K)	Remarks
300–600	Heat capacity nearly constant $C_v \simeq \frac{5}{2}R$
1000	Dissociation begins, C_v rises
5000	Dissociation virtually complete. System now consists of 2 moles of atomic hydrogen $C_v \simeq 3R$
10^4	Ionization of atoms noticeable, C_v rising
10^5	Ionization virtually complete. The system now consists of 2 moles of protons and 2 moles of electrons, C_v has leveled off at a value near $6R$
10^5–10^8	C_v remains nearly constant
10^9	Collisions of particles produce electron–positron pairs, C_v rising
10^{11}	Proton–proton collisions produce pions, C_v rising
10^{12}	Pions abundant. Energetic proton–proton collisions produce proton–antiproton pairs, C_v rising
$\simeq 1.2 \times 10^{12}$	Temperature has reached the ultimate limit. Addition of still more energy merely increases the number and variety of particles, C_v rising

How are we to interpret these results? An important point to note about the "data" is that increases in the heat capacity accompany events which increase the total number of particles (dissociation, ionization, pair-production, and so on). On the other hand C_v remained constant over temperature ranges where the number of particles did not change appreciably.

Since we prescribed that the experiment proceeds at constant volume there is no PdV work; the energy added to the system shows up as an increase of internal energy

$$\text{energy added} = \Delta U$$

The definition of C_v is

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V$$

Over a temperature range where C_v remains constant we may conclude that all of the energy added goes toward increasing the kinetic energies of the particles and this registers as a temperature increase

$$\text{energy added} = \Delta U = C_v\, \Delta T, \qquad C_v = \text{constant}$$

An increase of C_v with T shows that energy is being used for purposes beyond the increase of kinetic energy. These other uses are obvious; it takes energy to pull apart a hydrogen molecule. Energy is required to tear the electron from a hydrogen atom. In the process of pair-production, energy is converted from one form (kinetic) into another (mass).

The change of C_v from $(5/2)\,R$ to $3R$ brought about by molecular dissociation simply reflects the increase in the number of particles. The gas goes from 1 mole of diatomic molecules (each having 5 degrees of freedom) to 2 moles of monatomic atoms (each having 3 degrees of freedom). The values of $(5/2)\,R$ and $3R$ follow from the equipartition prescription of $\frac{1}{2}R$ per degree of freedom per mole. Ionization of the atomic hydrogen again doubles the number of particles. With protons and electrons alike possessing 3 degrees of freedom the heat capacity is again doubled.

Energy used to break atomic bonds is not available to increase the temperature. Processes which soak up energy by creating additional particles act as thermostats—they act to prevent an increase in temperature. Let E denote the energy characteristic of the particle-producing reaction. For example, in the dissociation of molecular hydrogen, $E = 4.48$ eV. For the creation of an electron–positron pair, $E = 2m_e c^2 \simeq 1$ MeV, the rest mass energy of the pair. At temperature T the average kinetic energy per particle is on the order of kT. The relation

$$kT \simeq \tfrac{1}{10}E \tag{18.67}$$

furnishes a *rough estimate* of the ignition temperature of the reaction. This rule-of-thumb follows from two considerations. The first point is an obvious one; particle-producing reactions are a result of collisions in which kinetic energy is converted into other forms. The second point is that a small but significant fraction of the particles will have kinetic energies at least 10 times as great as the average energy. Collisions between these above average particles initiate the reaction.

The relation (18.67) gives only a rough estimate—it ignores the fact that the pressure also influences the reaction rate. The greater the pressure, the more likely it is for particles to interact.

....

18.6.1 Using the rule-of-thumb (18.67), evaluate ignition temperatures for the following reactions:

(a) dissociation of H_2,
(b) ionization of H ($E_{ion} = 13.6$ eV),
(c) electron–positron pair production.

How do your values compare with the temperatures listed in the thought experiment?

....

When the temperature reaches the 10^8–10^9 K range, the production of electron–positron pairs becomes an important process. Electron–proton collisions and photon–proton collisions provide two mechanisms for electron–positron pair production.

At temperatures above 10^{11} K proton–proton collisions begin to produce *pions*.† The pions are the least massive members of the family of particles known as *hadrons*. Hadron is a generic name assigned to particles which interact through the powerful, short-range, nuclear force. In addition to the pions, the hadron family includes the neutron, the proton, and other more massive particles.

Experimentally, the number of hadrons of mass less than m is a rapidly increasing function of m. Theoretical considerations suggest that the number of hadrons with mass less than m increases exponentially for large m. This exponential increase in the number of mass states is countered by the Boltzmann factor $e^{-\beta\varepsilon}$ which governs their occupation probability. The ultimate temperature is a consequence of these two competing exponential factors.As the ultimate temperature is approached, energy added to the system is used to produce particles rather than to increase the kinetic energies of the existing particles. The thermostat action of particle production becomes complete—there can be no further temperature increase.

We have so far envisioned the system as a collection of particles which started out as hydrogen molecules and evolved into a "super-zoo" of hadrons as we added energy. On the other hand, a value for the ultimate temperature T_u is established by fitting the theory to experimental data on proton–proton collisions. In matching theory and experiment one assumes that two colliding protons momentarily form a system which achieves internal thermodynamic equilibrium. The two protons lose their identities and become a fireball—a

† There are three varieties of pions, π^0, π^+, and π^-. The superscript denotes the electric charge of the pion. The neutral pion (π^0) has a rest mass energy of 135 MeV. The charged pions $\pi^\pm$ have rest mass energies of 139.6 MeV.

system of many hadrons which distribute themselves over the available energy states in accordance with the laws of statistical mechanics.

After a lifetime of about 10^{-23} sec the fireball breaks up. In some instances two protons emerge, in which the case the event is described as an elastic scattering. However, as the kinetic energies of the colliding protons is increased the fireball debris becomes more varied and is likely to include other hadrons. Pion production is a frequent result. Two reactions which produce pions are

$$p + p \rightarrow p + n + \pi^+$$

and

$$p + p \rightarrow p + p + \pi^0$$

The production of proton–antiproton pairs becomes possible when the kinetic energies of the colliding protons reach the GeV range (1 GeV = 10^9 eV),

$$p + p \rightarrow p + p + p + \tilde{p}$$

The collision products are restricted only by the various conservation laws. For example, the conservation of electric charge prohibits the creation of a proton, but allows the production of proton–antiproton pairs.

....

18.6.2 The rest mass energy of a proton is 938 MeV. Two protons with equal kinetic energies collide head on, and produce a proton–antiproton pair

$$p + p \rightarrow p + p + p + \tilde{p}$$

What is the minimum kinetic energy required of the protons? Use (18.67) to estimate the ignition temperature for this creation.

....

At high energies the exponential character of the hadron mass spectrum makes itself felt. As the fireball temperature approaches the ultimate value it becomes more efficient to produce particles than to raise the temperature. The experimental evidence for thermodynamic equilibrium at an ultimate temperature is drawn from data on the momentum distribution of fireball fragments. If the fireball hadrons are in mutual thermodynamic equilibrium, they will have an isotropic distribution of momenta (in the rest frame of the fireball). One can measure the momentum distribution of the fragments in the plane perpendicular to the direction of motion of the fireball. Theory predicts that the distribution function for the perpendicular component of the momentum ($p_\perp$) is proportional to $\exp(-p_\perp c/kT_u)$, where T_u is the ultimate temperature. The fireball reactions include elastic scattering as well as the more exotic particle production schemes. Measurements of the momentum distribution of

elastically scattered protons is consistent with the expected form provided one takes $kT_u \simeq 160$ MeV or

$$T_u \simeq 1.2 \times 10^{12}\,\mathrm{K} \tag{18.68}$$

Can we understand why T_u has this particular value? The answer is "yes," and the argument is as follows: We observe, for the final time, that kT measures the average kinetic energy. We also note that the pion is the least massive hadron. The average kinetic energy cannot be raised above $m_\pi c^2$, the rest mass energy of the pion. Why? Because collisions very effectively convert kinetic energy into mass. It is not possible to increase the kinetic energies when pion-producing collisions soak up energy added to the system. The production of pions acts as a thermostat—as a governor for the temperature. We are therefore led to write

$$kT_u \simeq m_\pi c^2 \tag{18.69}$$

as the condition establishing the value of the ultimate temperature. With $m_\pi c^2 = 135$ MeV we find from (18.69) that $T_u \simeq 10^{12}$ K.

In brief we can say that the particular value of the ultimate temperature stems from the fact that particles with kinetic energies in excess of $m_\pi c^2$ will interact and produce pions. The existence of an upper limit for temperature seems destined to become the fourth law of thermodynamics.

REFERENCES

A survey of optical methods of temperature measurement, which draws heavily on the blackbody theory, is given in:

G. A. Hornbeck, Optical Methods of Temperature Measurement, *Appl. Optics* **5**, 179 (1966).

The essays by G. Gamow (Big Bang) and F. Hoyle (Continuous Creation) are included in:

J. R. Newman (ed.), "The Universe." Simon and Schuster, New York (1957).

A highly readable account of the remnant radiation is presented in:

P. J. E. Peebles and D. T. Wilkinson, The Primeval Fireball, *Sci. Am.* **216**, 28 (June 1967).

One measure of the scientific and technological impact of the laser is the number of Scientific American articles devoted to it! Included among these are the following:

A. L. Schawlow, Optical Masers, *Sci. Am.* **204**, 52 (June, 1961).

A. L. Schawlow, Advances in Optical Masers, *Sci. Am.* **209**, 34 (July 1963).

G. C. Pimentel, Chemical Lasers, *Sci. Am.* **215**, 32 (April 1966).

C. K. N. Patel, High Power Carbon Dioxide Lasers, *Sci. Am.* **219**, 23 (Aug. 1968).

A. L. Schawlow, Laser Light, *Sci. Am.* **219**, 120 (Sept. 1968).

D. R. Herriott, Applications of Laser Light, *Sci. Am.* **219**, 140 (Sept. 1968).

J. E. Faller and E. J. Wampler, The Lunar Laser Reflector, *Sci. Am.* **222**, 38 (March 1970).

M. J. Lubin and A. P. Fraas, Fusion by Laser, *Sci. Am.* **224**, 21 (June 1971).

The concept of a negative absolute temperature is discussed in:

M. W. Zemansky, "Temperatures Very High and Very Low." Van Nostrand-Reinhold, Princeton, New Jersey (1964).

The experiment which led to consideration of a negative temperature is described in the brief letter:

E. M. Purcell and R. V. Pound, A Nuclear Spin System at Negative Temperature, *Phys. Rev.* **81**, 279 (1951).

Subsequent studies of negative temperatures include:

N. F. Ramsey, Thermodynamics and Statistical Mechanics at Negative Absolute Temperatures, *Phys. Rev.* **103**, 20 (1956).

and

M. J. Klein, Negative Absolute Temperatures, *Phys. Rev.* **104**, 589 (1956).

The use of dimensionless ratios in astrophysics is discussed at length in:

E. E. Salpeter, Dimensionless Ratios and Stellar Structure, R. Marshak (ed.), "Perspectives in Modern Physics." Wiley (Interscience), New York (1966).

The black hole concept is explained in:

R. Ruffini and J. A. Wheeler, Introducing the Black Hole, *Phys. Today* **24**, 30 (Jan. 1971)

and

R. Penrose, Black Holes, *Sci. Am.* **226**, 38 (May 1972).

An incisive look at gravitational collapse is presented in:

K. S. Thorne, Gravitational Collapse, *Sci. Am.* **217**, 88 (Nov. 1967).

A broad view of cosmological aspects of gravitational collapse and "missing mass" is presented in:

J. A. Wheeler, Our Universe: The Known and the Unknown, *Am. Sci.* **56** (1), **1** (1968).

A broader view of cosmological questions is presented in the excellent monograph:

D. W. Sciama, "Modern Cosmology." Cambridge Univ. Press, London and New York (1971).

An informative and well-written article describing the structure of degenerate stars is:

M. Ruderman, Solid Stars, *Sci. Am.* **224**, 24 (Feb. 1971).

Additional arguments favoring the rotating neutron star model for pulsars is presented in:

J. P. Ostriker, The Nature of Pulsars, *Sci. Am.* **224**, 48 (Jan. 1971).

The phase separation in ^{3}He–^{4}He mixtures was detected using a magnetic resonance technique:

G. K. Walters and W. M. Fairbank, Phase Separation in He^3–He^4 Solutions. *Phys. Rev.* **103**, 262 (1956).

An excellent treatment of the theory and design of dilution refrigerators is presented in:

D. S. Betts, Helium Isotope Refrigeration, *Contemp. Phys.* **9**, 97 (1968).

A nontechnical discussion of dilution refrigeration and other quantum coolers is given in:

O. V. Lounasmaa, New Methods for Approaching Absolute Zero, *Sci. Am.* **221**, 26 (Dec. 1969).

The dependence of E_{34} on the ^{3}He fraction X is reported in:

D. O. Edwards, D. F. Brewer, P. Seligman, M. Skertic, and M. Yaqub, Solubility of He^3 in Liquid He^4 At 0 °K. *Phys. Rev. Letters* **15**, 773 (1965).

A comprehensive study of ^{3}He–^{4}He mixtures is presented in:

J. C. Wheatley, Dilute Solutions of ^{3}He in ^{4}He at Low Temperatures, *Am. J. Phys.* **36**, 181 (1968).

A selection of reprints, which includes Onnes' report of the discovery of superconductivity, has been published for the American Association of Physics Teachers. It is entitled:

"Superconductivity, Selected Reprints." Am. Inst. Phys., New York.

A splendid survey of all experimental aspects of superconductivity—up to 1955—is given in:

B. Serin, Superconductivity: Experimental Part, *Handbuch Phys.* **15**, 210 (1956).

A brilliant review of the theory of superconductivity is offered in:

H. Fröhlich, The Theory of the Superconductive State, *Reports Progress Phys.* **24**, 1 (1961).

Most parts of this review can be read with profit by an advanced student or an instructor. Sections 1 and 6 should be required reading for all aspiring scientists.

An inexpensive cryogenic system and an experiment suited to the measurement of transition temperatures above 2 K is described in:

E. P. Stillwell, Jr., R. L. Gardner, and H. T. Littlejohn, Experiments in Cryogenics for the Undergraduate, *Am. J. Phys.* **35**, 502 (1967).

The effects of ultra-high pressure on the superconducting properties of materials is surveyed in:

N. B. Brandt and N. I. Ginzburg, Superconductivity at High Pressure, *Sci. Am.* **224**, 83 (April 1971).

A recent review of progress toward the 30 K transition temperature is given in:

B. T. Matthias, The Search for High-Temperature Superconductors, *Phys. Today* **24**, 21 (Aug. 1971).

Type II superconductors are discussed in:

A. C. Rose-Innes, The New Superconductors, *Contemp. Phys.* **7**, 135 (1965).

Independent experiments confirming the existence of flux quantization were reported in:

B. S. Deaver, Jr., and W. M. Fairbank, Experimental Evidence For Quantized Flux in Superconducting Cylinders. *Phys. Rev. Letters* **7**, 43 (1961).

R. Doll and M. Nabauer, Experimental Proof of Magnetic Flux Quantization in A Superconducting Ring. *Phys. Rev. Letters* **7**, 51 (1961).

The quantized vortex pattern in superconductors has been studied directly using a novel photographic technique. See:

U. Essmann and H. Trauble, The Magnetic Structure of Superconductors, *Sci. Am.* **224**, 75 (March 1971).

See also:

P. J. Collins and J. E. Gordon, An Undergraduate Experiment Demonstrating Flux Quantization and Superconductivity, *Am. J. Phys.* **37**, 293 (1969).

Introductory level books dealing with elementary particles which may prove helpful are the following:

C. E. Swartz, "The Fundamental Particles." Addison-Wesley, Reading, Massachusetts (1965).

Chapter 6 focuses on the conservation laws.

K. W. Ford, "The World of Elementary Particles." Ginn Blaisdell, Boston, Massachusetts (1963).

Chapter 4 deals with the conservation laws.

Appendix A

Gaussian Integrals

The function $e^{-\alpha x^2}$ occurs in many areas of statistical analysis. We refer to $e^{-\alpha x^2}$ as a *Gaussian function*. The Maxwellian distribution function

$$f(v) = \text{constant} \cdot \exp\left(\frac{-\frac{1}{2}mv^2}{kT}\right)$$

is a Gaussian function. In kinetic theory it is frequently necessary to evaluate integrals involving the Gaussian function. Specifically, one needs to evaluate integrals having the form

$$G_n(\alpha) = \int_0^\infty x^n e^{-\alpha x^2}\, dx \tag{A.1}$$

The $G_n(\alpha)$ are termed *Gaussian integrals*. Note that although $G_n(\alpha)$ is defined by a definite integral its value depends on the parameter α. We can treat $G_n(\alpha)$ as a function of α. In particular we can *differentiate* $G_n(\alpha)$ with respect to α. By differentiating $G_n(\alpha)$ we obtain what is known as a *recurrence relation*. The recurrence relation for Gaussian integrals lets us compute $G_{n+2}(\alpha)$ from $G_n(\alpha)$, via differentiation. In kinetic theory one is generally interested in cases where n is an integer. If we know $G_0(\alpha)$, the recurrence relation readily yields $G_2(\alpha)$, which in turn gives $G_4(\alpha)$, and so on. From $G_1(\alpha)$ we can compute $G_3(\alpha)$, $G_5(\alpha)$, and so on. Thus only $G_0(\alpha)$ and $G_1(\alpha)$ need be evaluated directly; the recurrence relation takes over from there.

To derive the recurrence relation we differentiate $G_n(\alpha)$ with respect to α. The integral defining $G_n(\alpha)$ is a definite integral—its limits are constants, independent of α. Consequently the integration over x in (A.1) and a differentiation with respect to α may be interchanged. Thus

$$\frac{dG_n}{d\alpha} = \int_0^\infty x^n \frac{d}{d\alpha}(e^{-\alpha x^2})\,dx$$

Remembering that we differentiate with respect to α,

$$\frac{d}{d\alpha}(e^{-\alpha x^2}) = -x^2 e^{-\alpha x^2}$$

so that

$$\frac{dG_n}{d\alpha} = -\int_0^\infty x^{n+2} e^{-\alpha x^2}\,dx$$

The right side is $-G_{n+2}(\alpha)$ so that the recurrence relation is

$$G_{n+2}(\alpha) = -\frac{dG_n}{d\alpha} \tag{A.2}$$

As remarked earlier, only $G_0(\alpha)$ and $G_1(\alpha)$ need be evaluated directly. Consider first

$$G_1(\alpha) = \int_0^\infty x e^{-\alpha x^2}\,dx \tag{A.3}$$

The substitution

$$\alpha x^2 = z, \qquad 2\alpha x\,dx = dz$$

converts the right side to

$$G_1(\alpha) = \frac{1}{2\alpha}\int_0^\infty e^{-z}\,dz \tag{A.4}$$

However,

$$\int_0^\infty e^{-z}\,dz = -e^{-z}\Big]_0^\infty = 1$$

so that (A.4) gives

$$G_1(\alpha) = \tfrac{1}{2}\alpha^{-1} \tag{A.5}$$

Using the recurrence relation (A.2) we find

$$G_3(\alpha) = -\frac{dG_1}{d\alpha} = \frac{1}{2}\alpha^{-2}$$

$$G_5(\alpha) = -\frac{dG_3}{d\alpha} = \alpha^{-3} \tag{A.6}$$

$$G_7(\alpha) = -\frac{dG_5}{d\alpha} = 3\alpha^{-4}$$

The evaluation of $G_0(\alpha)$ is a bit more tricky. In fact, we compute $[G_0(\alpha)]^2$. The symmetry of the integrand allows us to write

$$G_0(\alpha) = \int_0^{\infty} e^{-\alpha x^2}\,dx = \tfrac{1}{2}\int_{-\infty}^{+\infty} e^{-\alpha x^2}\,dx$$

Furthermore, x is a dummy variable of integration so that we may write

$$G_0(\alpha) = \tfrac{1}{2}\int_{-\infty}^{+\infty} e^{-\alpha x^2}\,dx = \tfrac{1}{2}\int_{-\infty}^{+\infty} e^{-\alpha y^2}\,dy$$

and thus

$$[G_0(\alpha)]^2 = \tfrac{1}{4}\int_{-\infty}^{+\infty} e^{-\alpha x^2}\,dx \int_{-\infty}^{+\infty} e^{-\alpha y^2}\,dy$$

Grouping factors

$$[G_0(\alpha)]^2 = \tfrac{1}{4}\int_{-\infty}^{+\infty}\int e^{-\alpha(x^2+y^2)}\,dx\,dy \tag{A.7}$$

we can regard (A.7) as an integration over the entire x–y plane, with $dx\,dy$ denoting the differential element of area. By switching to polar coordinates we can immediately evaluate $[G_0(\alpha)]^2$. From Figure A.1 we see that $x^2+y^2 = r^2$.

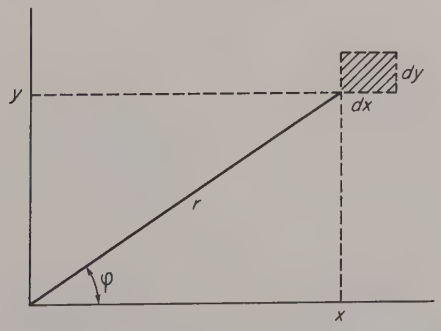


Figure A.1. Cartesian and polar coordinates. The hatched area = $dx\,dy$.

Thus

$$e^{-\alpha(x^2+y^2)} = e^{-\alpha r^2}$$

The integration

$$\int_{-\infty}^{+\infty}\int \cdots dx\, dy$$

is replaced by

$$\int_{\varphi=0}^{2\pi}\int_{r=0}^{\infty} \cdots r\, dr\, d\varphi$$

that is, $r\,dr\,d\varphi$ is the differential element of area in polar coordinates (see Figure A.2). The limits on φ and r are such as to cover the entire x–y plane.

We can now write

$$[G_0(\alpha)]^2 = \tfrac{1}{4}\int_0^{2\pi} d\varphi \int_0^{\infty} e^{-\alpha r^2} r\, dr \tag{A.8}$$

The φ integration gives 2π while

$$\int_0^{\infty} e^{-\alpha r^2} r\, dr = G_1(\alpha) = \tfrac{1}{2}\alpha^{-1}$$

Thus

$$[G_0(\alpha)]^2 = \frac{\pi}{4}\alpha^{-1}$$

which gives

$$G_0(\alpha) = \frac{\sqrt{\pi}}{2}\alpha^{-\frac{1}{2}} \tag{A.9}$$

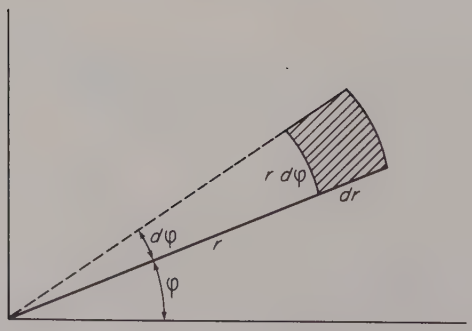


Figure A.2. Polar coordinate element of area. The hatched area = $r\,dr\,d\varphi$.

Using the recurrence relation we can generate $G_2(\alpha)$, $G_4(\alpha)$, and so on.

$$G_2(\alpha) = -\frac{dG_0}{d\alpha} = \frac{\sqrt{\pi}}{4}\alpha^{-3/2}$$

$$G_4(\alpha) = -\frac{dG_2}{d\alpha} = \frac{3\sqrt{\pi}}{8}\alpha^{-5/2}$$

$$G_6(\alpha) = -\frac{dG_4}{d\alpha} = \frac{15\sqrt{\pi}}{16}\alpha^{-7/2}$$

For ready reference the first eight Gaussian integrals are listed in Table A.1.

TABLE A.1

Gaussian Integrals

n	$G_n(\alpha) = \int_0^\infty x^n e^{-\alpha x^2}\,dx$	n	$G_n(\alpha) = \int_0^\infty x^n e^{-\alpha x^2}\,dx$
0	$\frac{1}{2}\sqrt{\pi}\alpha^{-1/2}$	4	$\frac{3}{8}\sqrt{\pi}\alpha^{-5/2}$
1	$\frac{1}{2}\alpha^{-1}$	5	α^{-3}
2	$\frac{1}{4}\sqrt{\pi}\alpha^{-3/2}$	6	$\frac{15}{16}\sqrt{\pi}\alpha^{-7/2}$
3	$\frac{1}{2}\alpha^{-2}$	7	$3\alpha^{-4}$

Appendix B

The Universal Character of β

In Chapter 15 we obtained a general relation for $\ln W$, with W being the thermodynamic probability

$$\ln W = N \ln Z(\beta) + \beta U \tag{15.53}$$

The parameter β originally entered as a convenience factor which rendered βU dimensionless. Considerations involving the internal energy of an ideal gas showed that

$$\beta = \frac{1}{kT} \tag{B.1}$$

at least for the ideal gas. We now establish the universal character of β, that is, we show that $\beta = 1/kT$ for all systems. The proof employs the "super system" device used in Section 15.9 to deduce the relationship between entropy and thermodynamic probability.

Let two systems, A and B, be in mutual thermodynamic equilibrium at a temperature T. Collectively A and B constitute a super system AB which is also in an equilibrium state at the temperature T. The populations of the phase-space cells of A and B are given by the Boltzmann distribution (15.33)

$$N_{i\mathrm{A}} = \frac{N_{\mathrm{A}} Q_{i\mathrm{A}}}{Z_{\mathrm{A}}(\beta_{\mathrm{A}})} \exp(-\beta_{\mathrm{A}} \varepsilon_{i\mathrm{A}}) \tag{B.2}$$

$$N_{k\mathrm{B}} = \frac{N_{\mathrm{B}} Q_{k\mathrm{B}}}{Z_{\mathrm{B}}(\beta_{\mathrm{B}})} \exp(-\beta_{\mathrm{B}} \varepsilon_{k\mathrm{B}}) \tag{B.3}$$

To establish the universal character of β we must show that $\beta_A = \beta_B$. Since A and B are arbitrary this step would show only that β has the same value for all systems. We have already established that $\beta = 1/kT$ for an ideal gas. Thus by demonstrating that $\beta_A = \beta_B$, then allowing B to be an ideal gas, we can show $\beta_A = 1/kT$, where A denotes any system. Because of the relation

$$W(\mathrm{AB}) = W(\mathrm{A}) \cdot W(\mathrm{B}) \tag{15.47}$$

we have

$$W(\mathrm{AB}) = \frac{N_A!}{\prod_{i=1}^{M_A} N_{iA}!} \frac{N_B!}{\prod_{k=1}^{M_B} N_{kB}!}$$

If one carries out the same maximization procedure for $\ln W(\mathrm{AB})$, namely,

$$\ln W(\mathrm{AB}) - \beta_s U_{AB}(\beta_s) = \ln W'(\mathrm{AB}) - \beta_s U'_{AB}(\beta_s)$$

the result is

$$N_{iA} N_{kB} = \frac{N_A Q_{iA} \exp(-\beta_s \varepsilon_{iA})}{Z_A(\beta_s)} \frac{N_B Q_{kB} \exp(-\beta_s \varepsilon_{kB})}{Z_B(\beta_s)} \tag{B.4}$$

From (B.2) and (B.3) we have

$$N_{iA} N_{kB} = \frac{N_A Q_{iA} \exp(-\beta_A \varepsilon_{iA})}{Z_A(\beta_A)} \frac{N_B Q_{kB} \exp(-\beta_B \varepsilon_{kB})}{Z_B(\beta_B)} \tag{B.5}$$

Comparing (B.4) and (B.5) leads to

$$\frac{Z_A(\beta_A)}{Z_A(\beta_s)} \exp[-(\beta_s - \beta_A)\varepsilon_{iA}] = \frac{Z_B(\beta_s)}{Z_B(\beta_B)} \exp[-(\beta_B - \beta_s)\varepsilon_{kB}] \tag{B.6}$$

On the left side stands a function referring to A alone; on the right side there stands a quantity referring to B alone. The only possible way for the equality to prevail is if both sides equal the same constant. This in turn is possible only if $\beta_s = \beta_A$; otherwise the exponential factor

$$\exp[-(\beta_s - \beta_A)\varepsilon_{iA}]$$

would be different for each different value of ε_{iA}. Likewise, the right side of (B.6) is a constant only if $\beta_s = \beta_B$. Thus we have

$$\beta_A = \beta_s, \qquad \beta_B = \beta_s$$

and so

$$\beta_A = \beta_B$$

Hence β has the same value for all systems at the same temperature. We have seen that $\beta = 1/kT$ for an ideal gas. It must therefore have this same value for all systems.

Subject Index

A

Absolute zero, 20, 123
  unattainability of, 124, 133
Activated complex, 312
Activation energy, 312
Adiabatic demagnetization, 176
Adiabatic process, 73
Angular momentum, 327
Antimatter, 42
Archimedes' principle, 13
Atomic size, 203
Avogadro's number, 7, 203

B

BCS theory, 398
Big bang theory, 347, 354
Black, J., 92
Blackbody radiation, 323, 347
Black hole, 365
Bohr, N., 327, 358
Boiling point, 17
Boltzmann, L., 49, 120
Boltzmann constant, 19
Boltzmann distribution, 287, 290
Boltzmann factor, 291
Bose–Einstein distribution, 341
Bose–Einstein statistics, 329, 340
Boson, 329

C

Caloric theory of heat, 43
Calorimetry, 45
Carnot, S., 79
Carnot cycle, 81
  entropy and, 102
Carnot engine, 86
Carnot's theorem, 87
Chandrasekhar, B. S., 372
Chandrasekhar limit, 366, 372
Chemical equilibrium, 8, 311
  condition for, 158
Chemical potential, 136, 154
  chemical equilibrium and, 158
  composition of neutron star and, 375
  of fermion gas, 335
  of ideal gas, 157
  phase equilibrium and, 157
Chemical reactions, 309
Clausius, R., 49, 119
Clausius–Clapeyron equation, 91, 167, 305

Coefficient
of linear expansion, 31, 33
of performance, 86
of volume expansion, 31, 125
Collision,
elastic, 205
star–star, 189
Compressibility, 34
Compton wavelength, 371
Cooper pair, 397
Critical field, in superconductor, 388, 398
Critical point, 93, 166
Curie's law, 68, 179, 316, 319
Cryogenics, 170

D

Dalton's law of partial pressures, 75
Degeneracy, quantum statistical, 342
Degenerate stars, 365
Degree of freedom, 130, 211
Density of states, 296
Deuterium, 42
Deuteron, 42
Differential, 23, 52
Dilution refrigerator, 376
Dissociation, 220, 310
Distribution,
Boltzmann, 287, 290
Bose–Einstein, 341
degenerate, 342
of distinguishable objects, 261
Fermi–Dirac, 333
of molecular speeds, 98, 196
most probable, 264, 274
Distribution function, 194
Maxwellian, 197, 234, 297
velocity space, 233

E

Effusion, 251
Einstein, A., 327, 358
Emission,
spontaneous, 358
stimulated, 358
Emissive power, monochromatic, 324
Energy, 37, 157
activation, 312
Coulomb, 310
dissociation, 309
equipartition of, 211, 215
Fermi, 335
interaction, 216
internal 48, 216, 225
potential, 225
quantization of, 326
Energy density, 222, 350
Energy spectrum, 282, 330
Enthalpy, 136, 147, 172, 302
Entropy, 102, 123
of blackbody radiation, 356
Carnot cycle and, 102
degrees of freedom and, 130
"frozen-in", 130
heat capacities and, 126
of ideal gas, 111, 115
as a measure of disorder, 107
statistical interpretation of, 297, 304
unit of, 106
Entropy formulation of second law, 119
Equation of state, 7, 19
of Curie paramagnet, 165
of ideal gas, 19, 72, 111, 203, 209
van der Waals, 225
virial, 225
Equipartition of energy, 211, 215

F

Factorial, 260
Fermi–Dirac distribution, 333
Fermi–Dirac statistics, 329
Fermi energy, 335
of electron gas, 337
Fermi surface, 337
Fermion, 329
First law of thermodynamics, 1, 37, 49, 53
adiabatic form, 50
equation form, 50, 52
magnetic version, 67
Fluctuation hypothesis, 120
Flux,
concept of particle, 248
expulsion of magnetic, 386, 398
internal energy, 253
magnetic, 66, 373, 399
Fluxoid, 399
Free energy,
Gibbs, 136, 151, 302
Helmholtz, 136, 145, 302

Free expansion, 111
Fusion,
  nuclear, 39
  thermonuclear, 40

G

Gamma function, 260
Gaussian function, 409
Gaussian integral, 409
Gibbs, J. W., 49
Gibbs free energy, 136, 151, 302
Gibbs' paradox, 111, 122

H

Hadron, 402
Heat,
  calorimetric definition of, 43
  of formation, 148
  geometric interpretation of, 116
  latent, 92, 162, 169, 308
  mechanical equivalent of, 50
  of reaction, 147
Heat capacity, 44, 69
  at constant pressure, 71, 218, 219
  at constant volume, 71, 217
  Debye, 70, 180
  electron gas, 339
  ideal gas, 72, 217
  magnetic, 178
Heat engine, 77
Heat exchanger, 174
Heat reservoir, 77
Helium,
  ^{4}He, 378
  ^{3}He, 378
Helmholtz free energy, 136, 145, 302
Hess's law, 150
Hydrogen bomb, 40, 53
Hydrogen burning, 365

I

Ideal gas, 19, 72, 111, 202
  entropy of, 111, 115
  internal energy of, 73, 216
  microscopic model of, 205
  partition function for, 292
  scale of temperature, 19
Internal energy, 50, 136
Inversion curve, 173
Inverted population, 361
Isotope effect, 395

J

Jeans, J., 327
Joule, J. P., 49, 53
Joule–Thomson coefficient, 144, 173
Joule–Thomson porous-plug experiment, 172

K

Kelvin, Lord, 7
Kelvin, unit of temperature, 16

L

Laplace, principle of insufficient reason, 189
Laser, 357, 360
Latent heat,
  of fusion, 92, 162
  of sumlimation, 92, 169
  for superconductor, 394
  of vaporization, 92, 96, 308
Law of thermodynamics,
  first, 1, 37, 49, 53
  second, 1, 77, 100
  third, 1, 123
  zeroth, 1, 9
Liquefaction, 170

M

Macrostate, 283
  most probable, 286, 300, 303
Magnetic cooling, 175
Magnetic field, 65
  critical, 388, 398
  of pulsar, 373
Magnetic flux, 66, 373, 399
Magnetic susceptibliity, 68, 175, 316
Magnetic temperature, 180
Magnetic work, 65, 175
Magneton, Bohr, 319
Mass–radius relation, 371
Maxwell, J. C., 49, 97
Maxwell's demon, 97, 101, 122
Maxwell relations, 125, 141
Maxwellian velocity distribution, 197, 234, 297
  experimental verification of, 243
  mean kinetic energy for, 239

Mayer, R., 49, 53, 73
Mean free path length, 200
Mean free time, 201
Mean value, 184
  relation to distribution function, 196, 235
Meissner effect, 386
Melting curve, 128
Method of mixtures, 46, 163
Microstate, 283
Momentum space, 280, 336
Monochromatic emissive power, 324

N

Nernst, W., 124
Neutron star, 365
Normalization condition, 185, 197, 233

O

Occupation probability, 334
Onnes, H. K., 384
Optical pumping, 361

P

Paramagnetism, Langevin theory of, 316
Partition function, 291
  relation to thermodynamic properties, 301
Path dependence,
  of heat, 51
  of work, 47, 60
Path of thermodynamic process, 55
Pauli exclusion principle, 330, 335
Penetration depth, 390
Perpetual motion machine, 38, 53, 79
Phase, 154, 161
  change, 154, 161, 167
  diagram, 165
  equilibrium, 157
  separation in ^{3}He–^{4}He, 378
Phase space, 278, 281, 329
  volume element of, 280, 281, 336
  volume of unit cell, 336, 346
Photon gas, 111, 221, 347, 356
Pion, 402
Planck, M., 323
Pressure, 10
  as a measure of energy density, 222
  melting, 128
  partial, 75
  units of, 12
Primeval fireball, 347, 355
Probability, 185
  a priori, 188
  conditional, 192
  distribution function, 195
  occupation, 334
  thermodynamic, 284, 303
Process,
  adiabatic, 73
  cyclic, 54
  irreversible, 55, 57
  quasi-static, 55
  reversible, 58
Pulsar, 366

Q

Quantization, in nature, 322
  of angular momentum, 328
  of energy, 323, 326
  of magnetic flux, 399
Quantum degeneracy, 342
Quantum statistical mechanics, 329
  reduction to classical statistics, 342
  relation to thermodynamics, 345

R

Reaction rate, chemical, 310
Refrigerator,
  Carnot, 86
  dilution, 376
Regelation, 168

S

Saddle-point integration, 265
Second law of thermodynamics, 1, 77, 100
  entropy formulation of, 119
  Kelvin–Clausius statements of, 78
Simon, F., 130, 135
Specific heat capacity, *see* Heat Capacity
Spherical coordinates, 236
Spin angular momentum, 327
Statistical mechanics, rationale of, 273
Steady-state theory, 354
Steam point, 17
Stefan–Boltzmann law, 324, 350
Stirling's approximation, 265, 269
Stirling engine, 89, 100, 101
Sublimation, 92, 169

Subsystem, 130
Superconductivity, 384
Superconductor,
  critical field for, 388, 398
  flux expulsion by, 386
  penetration depth for, 390
  transition temperature for, 385
  type I, 385
  type II, 385, 398
Susceptibility, magnetic, 68, 175

T

T–S diagrams, 116
Temperature,
  Debye, 70
  ideal gas, 19
  ignition, 401
  magnetic, 180
  mechanical interpretation of, 209
  scales of, 15, 20, 87
  thermodynamic, 87
  transition, 385
  ultimate, 399, 403
Temperature gradient, 253
Ten-degree rule, 314
Thermal conductivity, 253
Thermochemistry, 147
Thermodynamic efficiency, 83, 123
Thermodynamic equilibrium,
  relation to most probable macrostate, 300
  requirements for, 8
Thermodynamic potentials, 136
Thermodynamic temperature, 87
Third law of thermodynamics, 1, 123
Thomson, W., 7
Throttling process, 172
Torr, 12
Torricelli, 13
Transport of molecular properties, 247
Triple line, 166
Triple point, 16, 166

U

Ultimate temperature, 399, 403

V

Vapor pressure, 91, 305
Velocity space, 232
  distribution function, 233
  spherical, 236
Virial coefficients, 225
Virial equation of state, 225

W

White dwarf, 365
Wien's displacement law, 353
Work,
  geometrical interpretation of, 60
  magnetic, 65, 177
  path dependence of, 46, 59
  thermodynamic, 58

Z

Zeroth law of thermodynamics, 1, 9

Useful physical constants

Quantity	*Symbol*	*Value*
Gas constant	R	8.32×10^{7} erg/K
		1.99 calorie/K
Avogadro's number	N_0	6.02×10^{23}
Boltzmann constant	k	1.38×10^{-16} erg/K
Molar volume (gas at STP)		2.24×10^{4} cm^3
Mechanical equivalent of heat	J	4.19 joule/calorie
Planck's constant	h	6.63×10^{-27} erg-sec
Planck's constant divided by 2π ("h cross")	$\hbar$	1.05×10^{-27} erg-sec
Electronic charge	e	4.80×10^{-10} esu
Speed of light	c	3.00×10^{-10} cm/sec
Electron mass		0.911×10^{-27} gm
Proton mass		1.67×10^{-24} gm
Bohr magneton	μ_B	9.22×10^{-21} erg/gauss
Stefan-Boltzmann constant	σ	5.67×10^{-5} erg/sec-cm^2-K^4
Blackbody energy density constant	a	7.56×10^{-15} erg/cm^3-K^4
Gravitational constant	G	6.67×10^{-8} dyne-cm^2/gm^2
Bohr radius	a_0	5.29×10^{-9} cm

334844

QC
311.15
K44

ISBN 0-12-404050-0